

#4

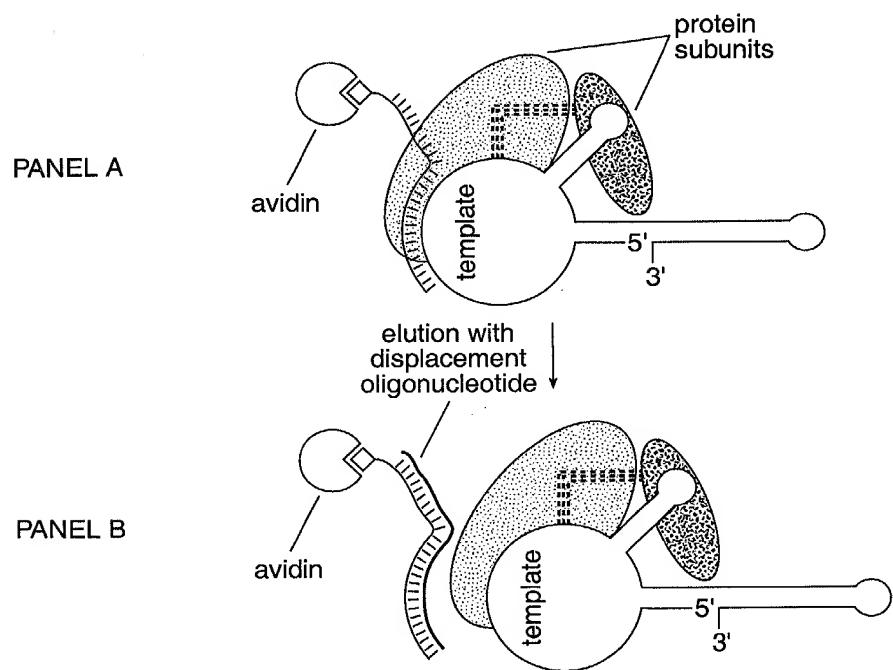
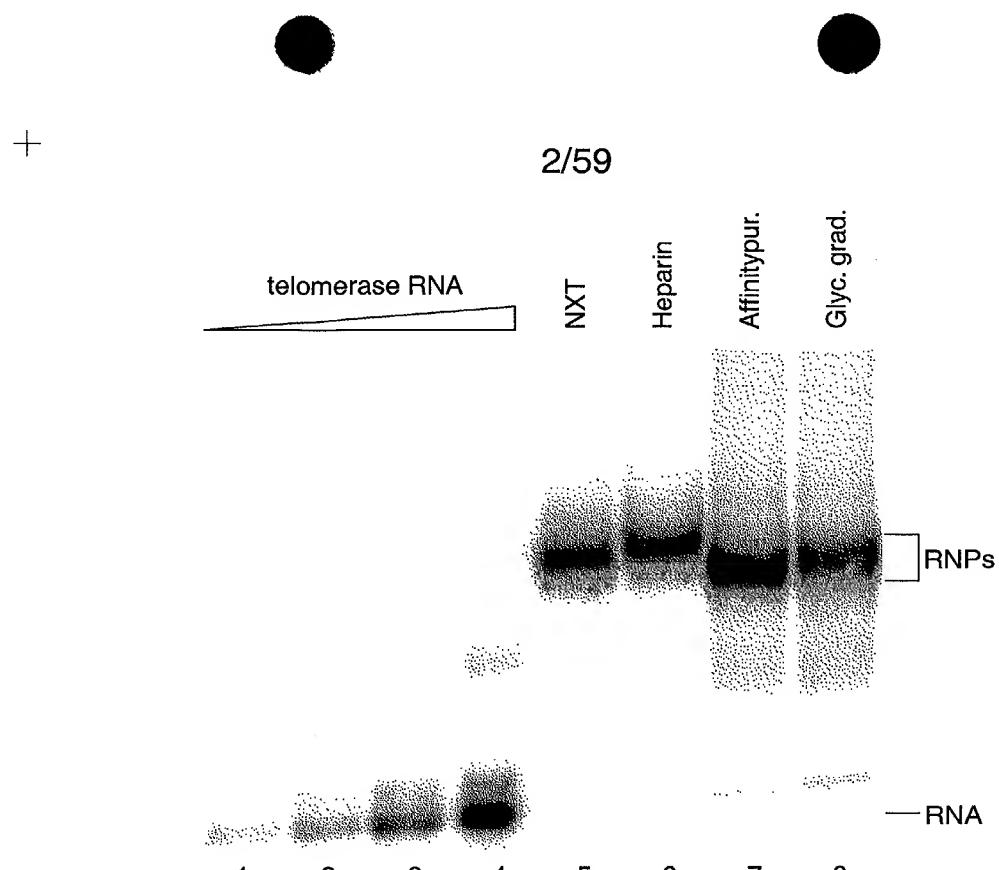
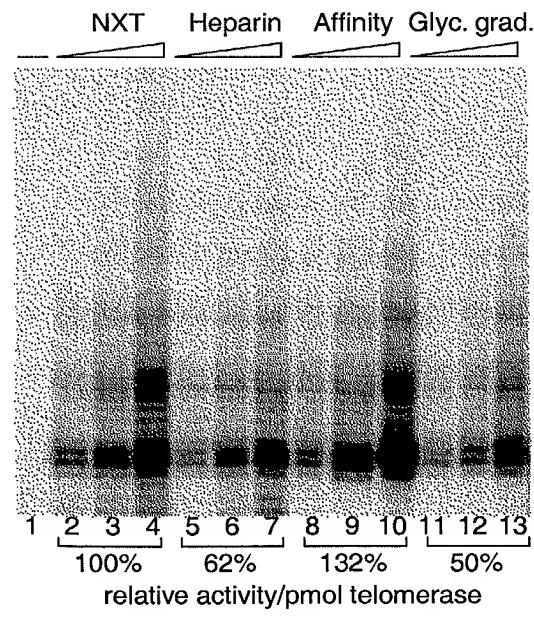


FIG. 1

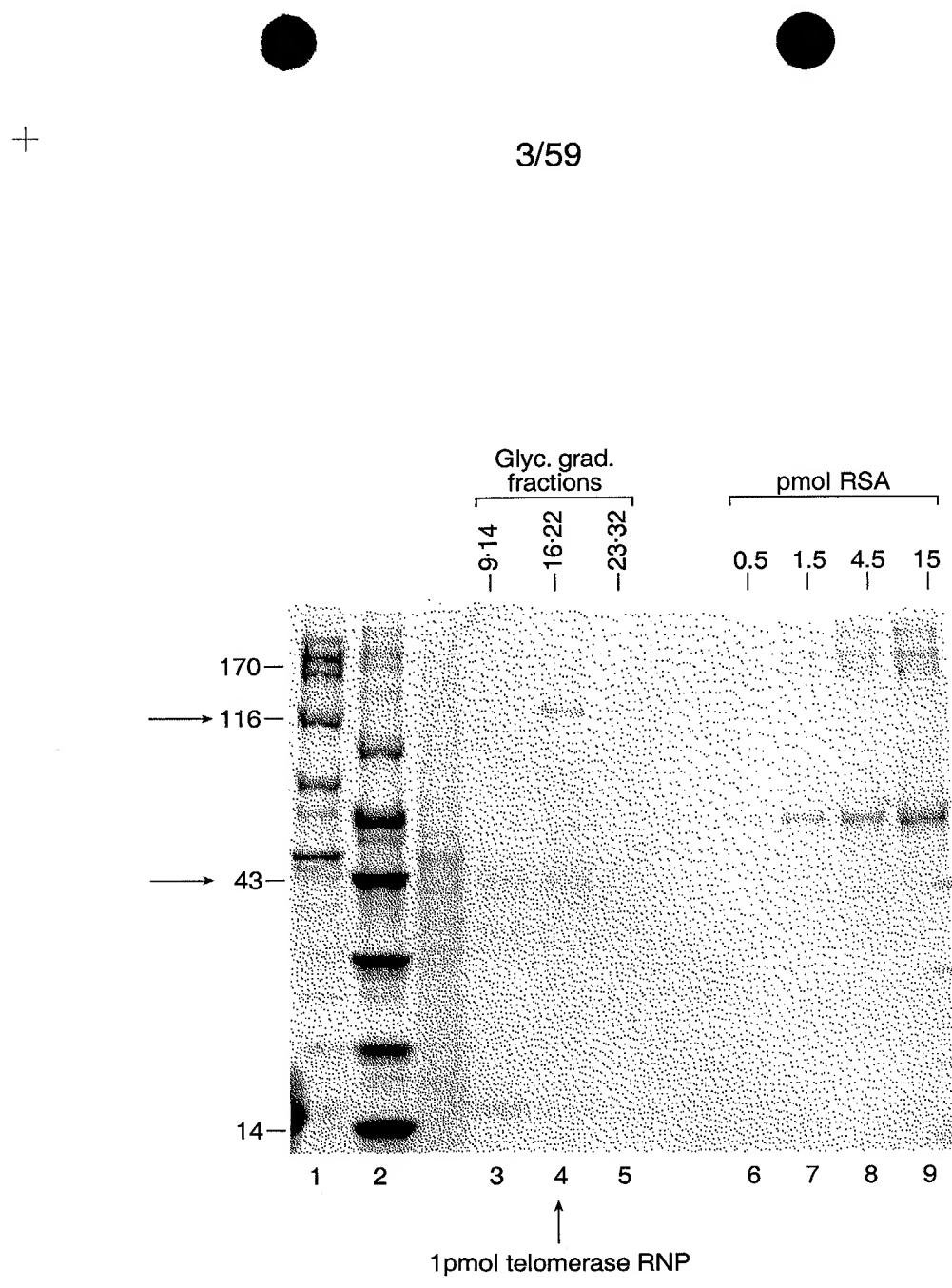


*FIG. 2*



*FIG. 3*

+



*FIG. 4*

+

+

4/59

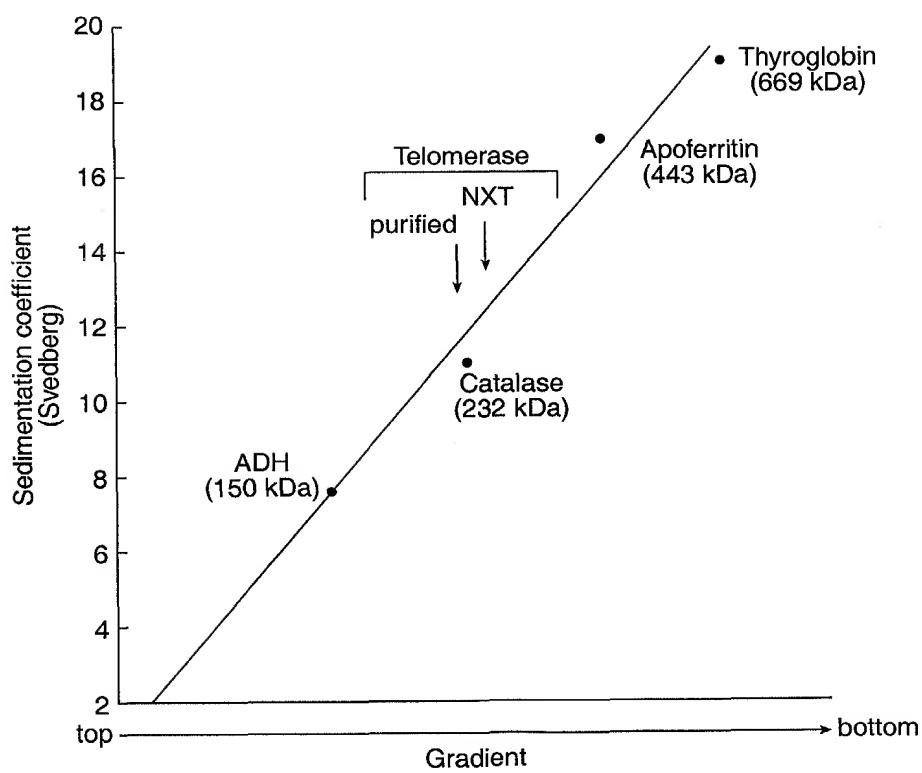


FIG. 5

+

5/59

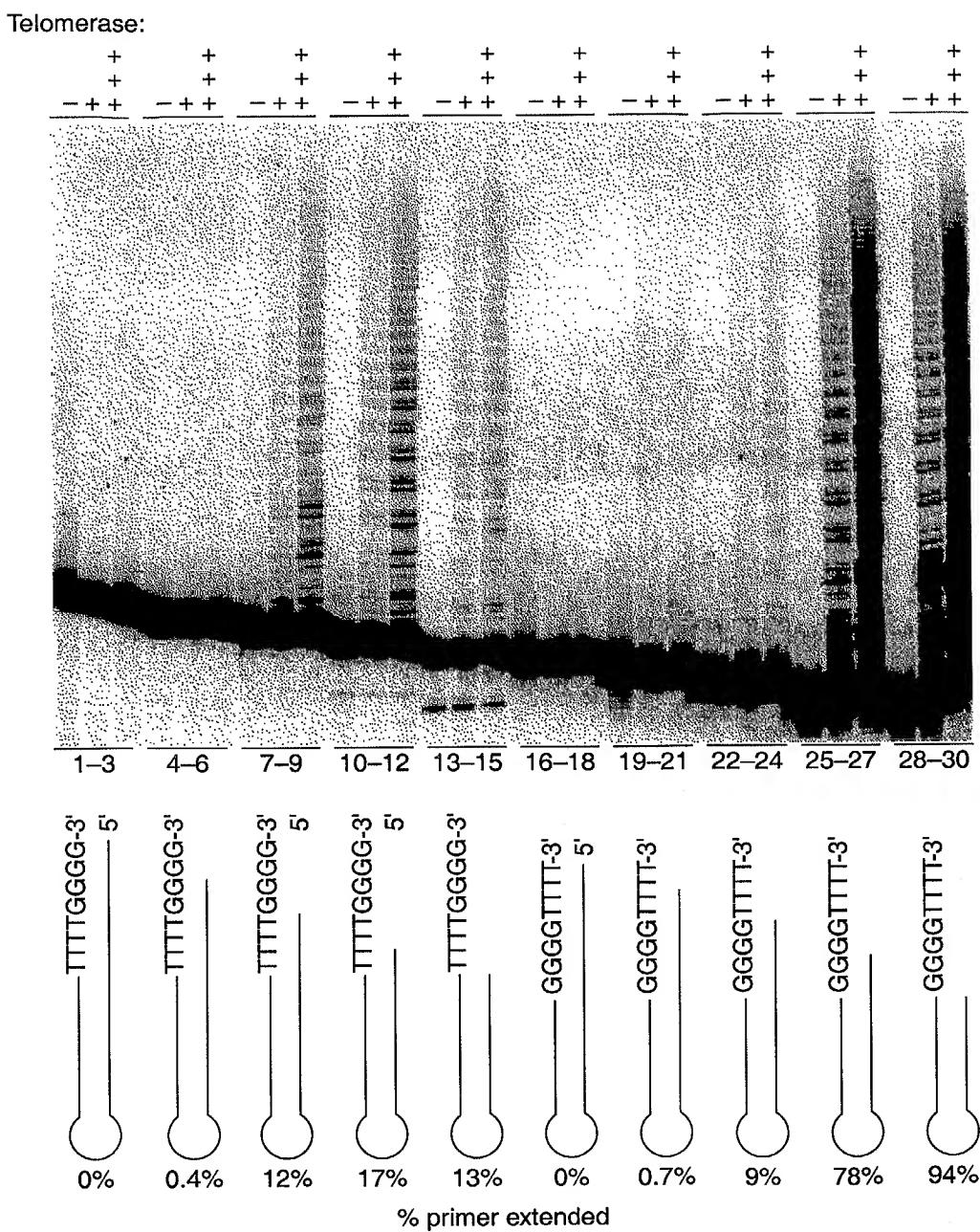


FIG. 6

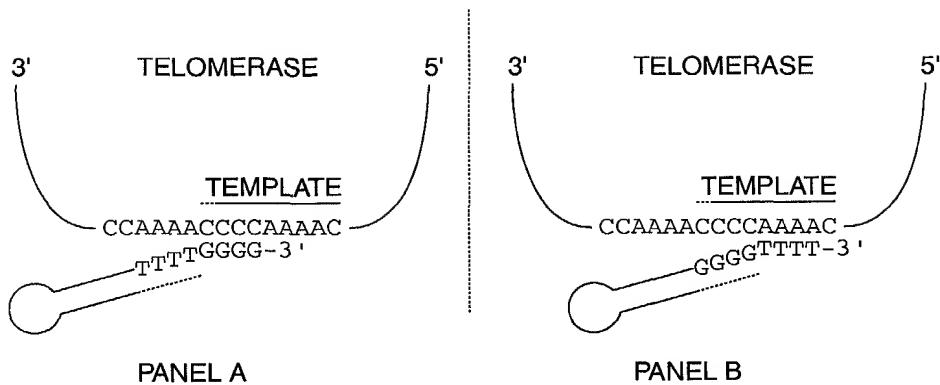


FIG. 7

1	CCCCAAAACC	CCAAAAACCC	AAAACCCCTA	TAAAAAAAGA	AAAATTGAG
51	GTA GTT TAGA	AATAAAATAT	TATTCCCGCA	CAAATGGAGA	TGGATATTGA
101	TTTGGATGAT	ATAGAAAATT	TACTTCCTAA	TACATTCAAC	AAGTAGAGCA
151	GCTCTGTAG	TGACAAGAAA	GGATGCAAAA	CATTGAAATC	TGGCTGAAA
201	TGCCTCTCAT	TGACTATTCC	AAAGTGCAA	AAACAAATTAG	AGTCTACTT
251	CTCGGGATGCA	AATCTTATA	ACGATTCTTT	CTTGAGAAAA	TTAGTTTAA
301	AAAGCGGAGA	GCAAAGAGTA	GAATTGAAA	CATTACTAAT	GT TAAATAA
351	AATCAGGTAA	TGAGGATTAT	TCTATTTTT	AGATCACTTC	T TAAAGGAGCA
401	TTATGGAGAA	AATTACTTAA	TA CAAAAGG	TAACAGTTT	GGATTATTTC
451	CCTAGCCAAC	AATGATGAGT	ATATTAAATT	CATATGAGAA	TGAGTCAAAG
501	GATCTCGATA	CATCAGACTT	ACCAAAGACA	AACTCGTAT	AAAAGC CAAAG
551	AAAAGTTTG	ATAATCGAAC	AGCAGAAGAA	CTTATTGCA	TTACTATTTC
601	TATGGGTTT	ATTACAATTG	TTTTAGGTAT	CGACGGTGAA	CTCCCAGTC
651	TTGAGACAA	TGAAAAAGCT	GT TTACA ACT	GAAGGAATCG	CAGTTCTGAA
701	AGTTCTGATG	TGATGCCAT	TATTTTGTGA	ATTAATCTCA	AATATCTTAT
751	CTCAATTAA	TGGATAGCTA	TA GAAAACAA	CCAAATAAAC	CATGCAAGTT
801	TAATGGATA	TACGTAAAT	CC TTGGGAC	AAATGCAAC	TGAATT TATA
851	TTGGATTCTT	AAAGCATAGA	TACACAGAAAT	GCTTTAGAGA	CTGATT TAGC
901	TTACAACAGA	TTACCTGTT	TGATTACTCT	TGCTCATCTC	TTATATCTTT
951	AAAAGAAGCA	GGCGAAATGA	AAAGAAGACT	AAAGAAGAG	ATTCAAAAT
1001	TTGTGATTTC	TTCTGTAAAC	GGAAATTAAAC	ACAAGAATAT	TAGCAACGAA
1051	AAAAGAAGAAC	AGCTATCAC	ATCTGATTCT	TTAAAGATT	CAAAATATTCC
1101	AGGTAAAGAGA	GATACATTC	T TAAATTCA	TATATTATAG	TTTTCTATT
1151	CACAGCTGTT	ATTTCTTTT	ATCTTAACAA	TATTTTTGTA	TTAGCTGGAA
1201	GTAAAAAGTA	TCAAATAAGA	GAAGCGCTAG	ACTGAGGTAA	CTTAGCTTAT
1251	TCACATTCA	AGATCGACCT	TCATATATCC	AATACGATGA	TAAGGAAACA
1301	GCAGTCATCC	GT TTTAAAAA	TA GTGCTATG	AGGACTAAAT	TTTTAGAGTC
1351	AAGAAATGGA	GCGGAAATCT	TAATCCCCAA	GAATTGCGTC	GATATTGCAA
1401	AAGAATCGAA	CTCTTAATCT	TT CGTTAATA	AGTATTACCA	ATCTTGATTG
1451	ATTGAAGAGA	TTGACGGAGC	AACTGCACAG	AAGATCATT	AAGAAATAAA
1501	GTA ACTTTTA	TTAATTAGAG	AATAAACTAA	ATTACTAATA	TAGAGATCAG
1551	CGATCTTCAA	TTGACGAAAT	AAAAGCTGAA	CTAAAGTTAG	ACAATAAAA
1601	ATACAAACCT	TGGTCAAAAT	ATTGAGGAAG	GAAAAGAAGA	CCAGTTAGCA
1651	AAGAAAAAAA	TAAGGCAATA	AATAAAATGA	GTACAGAACT	GAAGAAATAA
1701	AAGATT TATT	TTTTTCAAT	ATTTATTGAA	AAGAGGGTT	TTGGGGTTTT
1751	GGGGTTTGG	GG			

*FIG. 11*

DIG F5' End Labels

+

7/59



FIG. 8

+

1	AAAACCCCAA	AACCCAAAAA	CCCCTTTAG	AGCCCTGCAG	TTGGAAATAT
51	AACCTCAGTA	TTAATAAGCT	CAGATTTAA	ATATTAATTA	CAAAACCTAA
101	ATGGAGGTTG	ATGTTGATAA	TCAAGCTGAT	AATCATGGCA	TTCACTCAGC
151	TCTTAAGACT	TGTGAAGAAA	TTAAGAAGC	AAAACGTTG	TACTCTTGGA
201	TCCAGAAAGT	TATTAGATGA	AGAAAATCAAT	CTCAAAGTC	TTATAAAGAT
251	TTAGAAGATA	TTAAAATATT	TGCGCAGACA	AATATTGTTG	CTACTCCACG
301	AGACTATAAT	GAAGAAGATT	TTAAAGTTAT	TGCAAGAAAA	GAAGTATTTT
351	CAACTGGACT	AATGATCGAA	CTTATTGACA	AATGCTTAGT	TGAACCTCTT
401	TCATCAAGCG	ATGTTTCAGA	TAGACAAAAA	CTTCATGAT	TTGGATTTC
451	ACTTAAGGG	AATCAATTAG	CAAAGACCCA	TTTATTAACA	GCTCTTCAA
501	CTCAAAAGCA	GTATTCTTT	CAAGACGAAT	GGAAACCAAGT	TAGAGCAATG
551	ATTGGAAATG	AGCTCTCCG	ACATCTCTAC	ACTAAATATT	TAATATTCCA
601	GCGAACTTCT	GAAGGAACTC	TTGTTCAATT	TTGCCGGAAAT	AACGTTTTG
651	ATCATTGAA	AGTCAACGAT	AAAGTTGACA	AAAAGCAAA	AGGTGGAGCA
701	GCAGACATGA	ATGAACCTCG	ATGTTGATCA	ACCTGCAAAT	ACAATGTCAA
751	GAATGAGAAA	GATCACTTTC	TCAACAAACAT	CAACGTGCCG	AATTGGAATA
801	ATATGAAATC	AAGAACAGA	ATATTTTATT	GCACTCATTT	TAATAGAAAT
851	AACCAATTCT	TCAAAAAGCA	TGAGTTTG	AGTAACAAAA	ACAATATTC
901	ACCGATGGAC	AGAGCTCAGA	CGATATTAC	GAATATATT	AGATTTAATA
951	GAATTAGAAA	GAAGCTAAAA	GATAAGGTTA	TCGAAAAAAT	TGCCTACATG
1001	CTTGAGAAAG	TCAAAAGATT	TAACTTCAAC	TACTATTAA	CAAATCTTG
1051	TCCTCTTCCA	AAAAATTGGC	GGGAACGGAA	ACAAAAAATC	AAAAACTTGA
1101	TAAATAAAAC	TAGAGAAGAA	AAAGTCGAAGT	ACTATGAAGA	GCTGTTTAC
1151	TACACAACTG	ATAATAAATG	CGTCACACAA	TTTATTAATG	AATTCTTCTA
1201	CAATATACTC	CCCCAAAGACT	TTTTGACTGG	AAGAAACCCT	AAGAATTTC
1251	AAAAGAAAGT	TAAGAAATAT	GTGGAACCTAA	ACAAGCATGA	ACTCATTCA
1301	AAAAACTTAT	TGCTTGAGAA	GATCAATACA	AGAGAAATAT	CATGGATGCA
1351	GGTTGAGACC	TCTGCAAAGC	ATTTTTATTA	TTTTGATCAC	GAAAACATCT
1401	ACGTCTTATG	GAATTGCTC	CGATGGATAT	TCGAGGATCT	CGTCGTC
1451	CTGATTAGAT	GATTCTTCTA	TGTCACCGAG	CAACAGAAAA	GTTACTCAA
1501	AACCTTATTAC	TACAGAAAGA	ATATTTGGG	CGTCATTATG	AAAATGTCAA
1551	TCGCAGACTT	AAAGAAGGAA	ACGCTTGCTG	AGGTCCAAGA	AAAAGAGGTT
1601	GAAGAATGGA	AAAAGTCGCT	TGGATTGCA	CCTGGAAAAC	TCAGACTAAT
1651	ACCGAAGAAA	ACTACTTTCC	GTCCAATTAT	GACTTCAAT	AAGAAGATTG
1701	TAAATTTCAGA	CCGGAAGACT	ACAAAATTAA	CTACAAATAC	GAAGTTATTG
1751	AACTCTCACT	TAATGCTTAA	GACATTGAAG	AATAGAATGT	TTAAAGATCC
1801	TTTTGGATTTC	GCTGTTTTTA	ACTATGATGA	TGTAATGAAA	AAGTATGAGG
1851	AGTTTGTGTT	CAAATGGAAG	CAAGTTGGAC	AACCAAAACT	CTTCTTGC
1901	ACTATGGATA	TCGAAAAGTG	ATATGATAGT	GTAAACAGAG	AAAAACTATC
1951	ACATTCCTA	AAAACACTA	ATTACTTT	TTCAGATTTC	TGGATTATGA
2001	CTGCACAAAT	TCTAAAGAGA	AAGATAACA	TAGTTATCGA	TTCGAAAAC
2051	TTTAGAAAGA	AAGAAATGAA	AGATTATTTT	AGACAGAAAT	TCCAGAAAGAT
2101	TGCACTTGA	GGAGGACAAT	ATCCAACCTT	ATTCACTGTT	CTTGAAAATG
2151	ACAAAATGA	CTTAAATGCA	AAGAAAACAT	TAATTGTTGA	AGCAAAGCAA
2201	AGAAATTATT	TTAAGAAAGA	TAACTTACTT	CAACCAGTCA	TTAATATTG
2251	CCAATATAAT	TACATTAAC	TTAATGGGAA	GTTTTATAAA	CAAACAAAAG
2301	GAATTCCCTCA	AGGTCTTGA	GTTCATCAA	TTTGTCTC	ATTTTATTAT
2351	GCAACATTAG	AGGAAAGCTC	CTTAGGATT	CTTAGAGATG	AATCAATGAA

FIG. 9

2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC  
 2451 TTTTGATTAC AACTCAAGAG ATAATGCAG TATTGTTTAT TGAGAAACTT  
 2501 ATAAACGTAA GTCGTGAAAA TGGATTTAAA TTCAATATGA AGAAACTACA  
 2551 GACTAGTTT CCATTAAGTC CAAGCAAATT TGAAAATAC GGAATGGATA  
 2601 GTGTTGAGGA GCAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC  
 2651 TCAATTGATA TGAAAACTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT  
 2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGAAACAA AAGAAAGCAT  
 2751 CAATGTGGCT CAAGAAGAAA CTAAGTCGT TTTAATGAA TAACATTACC  
 2801 CATTATTTA GAAAGACGAT TACAACCGAA GACTTTGCGA ATAAAACCT  
 2851 CAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAG  
 2901 AATACAAGGA CCACTTTAACG AAGAACTTAG CTATGAGCAG TATGATCGAC  
 2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA  
 3001 CCTTGTGTGC AATATTAAGG ATACAATTTC TGGAGAGGAG CATTATCCAG  
 3051 ACTTTTCCT TAGCACACTG AAGCACTTA TTGAAATATT CAGCACAAAAA  
 3101 AAGTACATT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA  
 3151 GCTAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA  
 3201 CTATTCTAAC TTATTTGGG AAGTTAATT TCAATTTTG TCTTATATAC  
 3251 TGGGGTTTGTG GGGTTTGGG GTTTTGGGG

**FIG. 9**  
(CONTINUED)

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQLVKIRC RNQSQSHYKD  
 51 LEDIKIFAQI NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL  
 101 SSDVSDRQK LQCFGQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM  
 151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA  
 201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTTR IFYCTHFNRN  
 251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DVIEKIAYM  
 301 LEVKVKDFNPN YYLTKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS  
 351 YTTDNKCVTQ FINEFFYNIL PKDFLTGRNR KNFQKKVKKY VELNKHELIH  
 401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKL RWIFEDLVVS  
 451 LIRCFFYVTE QQKSYSKTYY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV  
 501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL  
 551 NSHMLMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA  
 601 TMDIEKCYDS VNREKLSTFL KTTKLLSSDF WIMTAQILKR KNNIVIDSKN  
 651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKTLLIVEAKQ  
 701 RNYFKKDNLQ QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSIILSSFYY  
 751 ATLEESSSLGF LRDESMNPEN PVNVNLLMRLT DDYLLITTTQE NNAVLFIEKL  
 801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI  
 851 SIDMKTALM PNINLRIEGI LCTLNLNMQT KKASMWLKKK LKSFLMNNIT  
 901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID  
 951 LEVSKIIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK  
 1001 KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A

**FIG. 10**

10/59

1 CCCCCAAAACCCAAAACCCAAAACCCCTATAAAAAAGAAAAATTGAGGTAGTTAGA 60  
 GGGGTTTTGGGGTTTGGGTTGGGATATTTTTCTTTTAACTCATCAAATCT

a	P	Q	N	P	K	T	P	K	P	L	*	K	K	K	L	R	*	F	R	-
b	P	K	T	P	K	P	Q	N	P	Y	K	K	R	K	N	C	G	S	L	E
c	P	K	P	Q	N	P	K	T	P	I	K	K	E	K	I	E	V	V	*	K

61 AATAAAATATTATTCCCGCACAAATGGAGATGGATATTGATTGGATGATAGAAAATT 120  
 TTATTTATAATAAGGGCGTGTACCTCTACCTATAACTAAACCTACTATATCTTTAA

a	N	K	I	L	F	P	H	K	W	R	W	I	L	I	W	M	I	*	K	I	-
b	I	K	Y	Y	S	R	T	N	G	D	G	Y	C	F	G	C	Y	R	K	F	
c	*	N	I	I	P	A	Q	M	E	M	D	I	D	L	D	D	I	E	N	L	

121 TACTTCCTAACATCAACAAGTATAGCAGCTTGTAGTGACAAGAAAGGATGCAAA 180  
 ATGAAGGATTATGTAAGTTGTCATATCGTCGAGAACATCACTGTTCTTCCTACGTTT

a	Y	F	L	I	H	S	T	S	I	A	A	L	V	V	T	R	K	D	A	K	-
b	T	S	*	Y	I	Q	Q	V	*	Q	L	L	*	*	Q	E	R	M	Q	N	
c	L	P	N	T	F	N	K	Y	S	S	S	C	S	D	K	K	G	C	K	T	

181 CATTGAAATCTGGCTCGAAATCGCCTTCATTGACTATTCAAAGTTGCAAAAACAATTAG 240  
 GTAACTTAGACCGAGCTTAGCGGAAGTAACTGATAAGGTTCAACGTTTGTGTTAAC

a	H	C	N	L	A	R	N	R	L	H	C	L	F	Q	S	C	K	N	N	*	-
b	I	E	I	W	L	E	I	A	F	I	D	Y	S	K	V	A	K	T	I	R	
c	L	K	S	G	S	K	S	P	S	L	T	I	P	K	L	Q	K	Q	L	E	

241 AGTTCTACTTCTCGGATGCAAATCTTATAACGATTCTTCTTGAGAAAATTAGTTTAA 300  
 TCAAGATGAAGAGCCTACGTTAGAAATATTGCTAAGAAAGAACTCTTTAATCAAATT

a	S	S	T	S	R	M	Q	I	F	I	T	I	L	S	C	E	N	*	F	*	-
b	V	L	L	L	G	C	K	S	L	*	R	F	F	L	E	K	I	S	F	K	
c	F	Y	F	S	D	A	N	L	Y	N	D	S	F	L	R	K	L	V	L	K	

301 AAAGCGGAGAGCAAAGAGTAGAAACATTACTAATGTTAAATAAAATCAGGAA 360  
 TTTCGCCTCTCGTTCTCATCTTAACCTTGTAATGATTACAATTATTAGTCCATT

a	K	A	E	S	K	E	*	K	L	K	H	Y	*	C	L	N	K	I	R	*	-
b	K	R	R	A	K	S	R	N	C	N	I	T	N	V	*	I	K	S	G	N	
c	S	G	E	Q	R	V	E	I	E	T	L	L	M	F	K	*	N	Q	V	M	

361 TGAGGATTATTCTATTAGATCACTTCTTAAGGAGCATTATGGAGAAAATTACTTAA 420  
 ACTCCTAACATAGATAAAATCTAGTGAAAGAAATTCCCTCGTAATACCTCTTTAATGAATT

a	C	G	L	F	Y	F	L	D	H	F	L	R	S	I	M	E	K	I	T	*	-
b	E	D	Y	S	I	F	*	I	T	S	*	G	A	L	W	R	K	L	L	N	
c	R	I	I	L	F	F	R	S	L	L	K	E	H	Y	G	E	N	Y	L	I	

FIG. 12

11/59

421 TACTAAAAGGTAAACAGTTGGATTATTCCTAGCCAACAATGATGAGTATATTAAATT  
480 ATGATTTCCATTGTCAAACCTAATAAAGGGATCGGTTGTTACTACTCATATAATTAA

a Y \* K V N S L D Y F P S Q Q C C V Y \* I -  
b T K R \* T V W I I S L A N N D E Y I K F -  
c L K G K Q F G L F P \* P T M M S I L N S -

CATATGAGAATGAGTCAAAGGATCTGATACATCAGACTTACCAAAGACAAACTCGCTAT  
481 GTATACTCTTACTCAGTTCCCTAGAGCTATGTAGTCTGAATGGTTCTGTTGAGCGATA 540

a H M R M S Q R I S I H Q T Y Q R Q T R Y -  
b I C E C V K G S R Y I R L T K D K L A I -  
c Y E N E S K D L D T S D L P K T N S L \* -

AAAACGCAAGAAAAAGTTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTG  
541 600 TTTTGCCTTCTTTCAAACATTAGCTTGTCTCTGAATAACGTAATGATAAGC

a K T Q E K V C \* S N S R R T Y C I Y Y S -  
b K R K K K F D N R T A E E L I A F T I R -  
c N A R K S L I I E Q Q K N L L H L L F V -

TATGGGTTTATTACAATTGTTTAGGTATCGACCGTGAACCCGAGTCTTGAGACAAT  
601 660 ATACCCAAAATAATGTTAACAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA

a Y G F Y Y N C F R Y R R C T P E S C D N -  
b M G F I T I V L G I D G E L P S L E T I -  
c W V L L Q L F \* V S T V N S R V L R Q L -

TGAAAAAGCTGTTACAACATGAAGGAATCGCAGTTCTGAAAGTCTGATGTGTATGCCAT  
661 720 ACTTTTGCACAAATGTTGACTTCCTTAGCGTCAAGACTTCAAGACTACACATACGGTA

a C K S C L Q L K E S Q F C K F \* C V C H -  
b E K A V Y N C R N R S S E S S D V Y A I -  
c K K L F T T E G I A V L K V L M C M P L -

TATTTGTGAATTAATCTCAAATATCTTATCTCAATTAAATGGATAGCTATAGAAACAAA  
721 780 ATAAAACACTTAATTAGAGTTATAGAATAGAGTTAACCTATCGATATCTTGT

a Y F V N \* S Q I S Y L N L M D S Y R N K -  
b I L C I N L K Y L I S I \* W I A I E T N -  
c F C E L I S N I L S Q F N G \* L \* K Q T -

CCAAATAAACCATGCAAGTTAATGGAATATACGTTAACCTTGGGACAAATGCACAC  
781 840 GGTTTATTGGTACGTTCAAATTACCTTATATGCAATTAGGAAACCTGTTACGTGTG

a P N K P C K F N G I Y V K S F G T N A H -  
b Q I N H A S L M E Y T L N P L G Q M H T -  
c K \* T M Q V \* W N I R \* I L W D K C T L -

TGAATTATATTGGATTCTTAAAGCATAGATACACAGAATGCTTAGAGACTGATTTAGC  
841 900 ACTTAAATATAACCTAAGAATTGCTATCTATGTCTTACGAAATCTGACTAAATCG

a C I Y I G F L K H R Y T E C F R D C F S -  
b E F I L D S \* S I D T Q N A L E T D L A -  
c N L Y W I L K A \* I H R M L \* R L I \* L -

FIG. 12  
(CONTINUED)

12/59

901 TTACAACAGATTACCTGTTTGATTACTCTTGCTCATCTCTTATATCTTTAAAAGAAGCA  
 960 AATGTTGTCTAATGGACAAACTAATGAGAACGAGTAGAGAATATAGAAATTTCTTCGT  
 a L Q Q I T C F D Y S C S S L I S L K E A -  
 b Y N R L P V L I T L A H L L Y L \* K K Q -  
 c T T D Y L F C L L L I S Y I F K R S R -  
 GGCGAAATGAAAAGAAGACTAAAGAAAGAGATTCAAAATTTGTTGATTCTCTGTAACC  
 961 1020 CCGCTTTACTTTCTGATTTCTCTAAAGTTAAACAACTAAGAACAGACATTGG  
 a G E M K R R L K K E I S K F V D S S V T -  
 b A K C K E D \* R K R F Q N L L I L L \* P -  
 c R N E K K T K E R D F K I C C F F C N R -  
 GGAATTAACAACAAGAATATTAGCAACGAAAAGAAGAGACTATCACAACTCCTGATTC  
 1021 1080 CCTTAATTGTTCTTATAATCGTTGCTTTCTCTCGATAGTGTAGGACTAAG  
 a G I N N K N I S N E K E E E L S Q S C F -  
 b E L T T R I L A T K K K S Y H N P D S -  
 c N \* Q Q E Y \* Q R K R R R A I T I L I L -  
 TTAAAGATTTCAAAAATTCAGGTAAAGAGAGATACTTCATTAAATTCAATATTATAG  
 1081 1140 AATTCTAAAGTTTTAAGGCCATTCTCTATGTAAGTAATTAAAGTATATAATATC  
 a L K I S K I P G K R D T F I K I H I L \* -  
 b \* R F Q K F Q V R E I H S L K F I Y Y S -  
 c K D F K N S R \* E R Y I H \* N S Y I I V -  
 TTTTCATTTCACAGCTGTTATTTCTTTATCTTAACAAATTTTGATTAGCTGGAA  
 1141 1200 AAAAGTAAAGTGTGACAATAAAAGAAAATAGAATTGTTATAAAAACTAATCGACCTT  
 a F F I S Q L L F S F I L T I F F D \* L E -  
 b F S F H S C Y F L L S \* Q Y F L I S W K -  
 c F H F T A V I F F Y L N N I F C L A G S -  
 GTAAAAAGTATCAAATAAGAGAACGCGTAGACTGAGGTAACCTAGCTTATTCACATTCA  
 1201 1260 CATTTTCATAGTTATTCTCTCGCGATCTGACTCCATTGAATCGAATAAGTGTAAAGTA  
 a V K S I K \* E K R \* T E V T \* L I H I H -  
 b \* K V S N K R S A R L R \* L S L F T F I -  
 c K K Y Q I R E A L D C G N L A Y S H S \* -  
 AGATCGACCTTCATATATCCAATACGATGATAAGGAAACAGCAGTCATCCGTTTAAAAAA  
 1261 1320 TCTAGCTGGAAGTATAGTTATGCTACTATTCCCTTGTGTCAGTAGGAAAATTTT  
 a R S T F I Y P I R C \* G N S S H P F \* K -  
 b D R P S Y I Q Y D D K E T A V I R F K N -  
 c I D L H I S N T M I R K Q Q S S V L K I -  
 TAGTGCTATGAGGACTAAATTTAGAGTCAGAAATGGAGCCGAAATCTTAATCAAAAAA  
 1321 1380 ATCACGATACTCCTGATTAAATCTCAGTTACCTCGGCTTTAGAATTAGTTTT  
 a \* C Y E D \* I F R V K K W S R N L N Q K -  
 b S A M R T K F L E S R N G A E I L I K K -  
 c V L C G L N F \* S Q E M E P K S \* S K R -

FIG. 12  
(CONTINUED)

GAATTGCGTCGATATTGCAAAAGAATCGAACTCTAAATCTTCGTTAATAAGTATTACCA  
 1381 -----+-----+-----+-----+-----+-----+-----+-----+ 1440  
 CTTAACGCAGCTATAACGTTTCTTAGCTTGAGATTAGAAAGCAATTATTCATAATGGT

a	E	L	R	R	Y	C	K	R	I	E	L	*	I	F	R	*	*	V	L	P	-
b	N	C	V	D	I	A	K	E	S	N	S	K	S	F	V	N	K	Y	Y	Q	-
c	I	A	S	I	L	Q	K	N	R	T	L	N	L	S	L	I	S	I	T	N	-

ATCTTGATTGATTGAAGAGATTGACGAGGCAACTGCACAGAAGATCATTAAAGAAATAAA  
 1441 -----+-----+-----+-----+-----+-----+-----+-----+ 1500  
 TAGAAACTAACTAACTTCTCTAACTGCTCCGTTGACGTGTCTCTAGTAATTTCTTTATTT

a	I	L	I	D	C	R	D	*	R	G	N	C	T	E	D	H	*	R	N	K	-
b	S	C	L	I	E	E	I	D	E	A	T	A	Q	K	I	I	K	E	I	K	-
c	L	D	C	L	K	R	L	T	R	Q	L	H	R	R	S	L	K	K	*	S	-

GTAACTTTTATTAAATTAGAGAATAAAACTAAATTACTAATATAGAGATCAGCGATCTCAA  
 1501 -----+-----+-----+-----+-----+-----+-----+-----+ 1560  
 CATTGAAAATAATTAATCTCTTATTTGATTAAATGATTATATCTCTAGTCGCTAGAAGTT

a	V	T	F	I	N	*	R	I	N	*	I	T	N	I	E	I	S	D	L	Q	-
b	*	L	L	L	I	R	E	*	T	K	L	L	I	*	R	S	A	I	F	N	-
c	N	F	Y	*	L	E	N	K	L	N	Y	*	Y	R	D	Q	R	S	S	I	-

TTGACGAAATAAAAGCTGAACTAAAGTTAGACAATAAAAAACACCTTGGTCAAAAT  
 1561 -----+-----+-----+-----+-----+-----+-----+-----+ 1620  
 AACTGCTTTATTTCGACTTGATTCAATCTGTTATTTTATGTTGGAACCAGTTTA

a	L	T	K	*	K	L	N	*	S	*	T	I	K	N	T	N	L	G	Q	N	-
b	C	R	N	K	S	C	T	K	V	R	Q	*	K	I	Q	T	L	V	K	I	-
c	D	E	I	K	A	E	L	K	L	D	N	K	K	Y	K	P	W	S	K	Y	-

ATTGAGGAAGGAAAAGAAGACAGTTAGCAAAAGAAAAATAAGGCATAAAATAAAATGA  
 1621 -----+-----+-----+-----+-----+-----+-----+-----+ 1680  
 TAAACTCCTCCTTTCTCTGGTCAATCGTTTCTTTTATTCGTTATTTTACT

a	I	E	E	G	K	E	D	Q	L	A	K	E	K	I	R	Q	*	I	K	C	-
b	L	R	K	E	K	K	T	S	*	Q	K	K	K	*	G	N	K	*	N	E	-
c	C	G	R	K	R	R	P	V	S	K	R	K	N	K	A	I	N	K	M	S	-

GTACAGAAGTGAAGAAATAAAAGATTATTTTTCAATAATTATTGAAAAGAGGGGTT  
 1681 -----+-----+-----+-----+-----+-----+-----+-----+ 1740  
 CATGTCTTCACCTCTTATTTCTAAATAAAAAAGTTATTAAATAACTTTCTCCCCAA

a	V	Q	K	C	R	N	K	R	F	I	F	F	N	N	L	L	K	R	G	V	-
b	Y	R	S	E	E	I	K	D	L	F	F	S	I	I	Y	C	K	E	G	F	-
c	T	E	V	K	K	*	K	I	Y	F	F	Q	*	F	I	E	K	R	G	F	-

TTGGGGTTTGGGGTTTGGGG  
 1741 -----+-----+-----+-----+-----+-----+-----+-----+ 1762  
 AACCCCCAAAACCCCAAAACCC

a	L	G	F	W	G	F	G	-
b	W	G	F	G	V	L	G	-
c	G	V	L	G	F	W	-	

**FIG. 12**  
(CONTINUED)

*FIG. 13*

15/59

798	EKLINVSRENGFKFNMKK.LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD	846
	..::: . .   .. . :: .. ::::  . ....	
577	.....PGDEL.RPSMQKLLQEKGKLGGG..TDFPYECIDEWTKNKTSHVD	617
847	WIGISIDMKTLLALMPNIINLRLEGILCTLNLNMQTKKASMWLKKLKSFLM	896
	.   .. . :  :  . :  . :  . :  . :  . :  . :  . :  . :	
618	NIVILSDMMIAEGYSDINVRGSSIVNSI.....KKYKDEVN	653
897	NNITHYFRKTITTEDFANKTLNKLFIGGGYKYMCAKEYKD.HFKKNLAM	945
	. : .. :::  . :  . :  . :  . :  . :  . :  . :  . :  . :	
654	PNIKIF...AVDLEGYG.....KCLNLGDEFNNENNYIKIFGM	687
946	SSMIDLLEVSKIYISVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFIE	995
	..  . :  . :  . :  . :  . :  . :  . :  . :  . :  . :  . :	
688	SDSI.....LKFISAKQGGA.....NMVE	706
996	IFSTKKYIFNRVC 1008	
	:: .. ::::	
707	VI..KNFALOKIG 717	

**FIG. 13**  
(CONTINUED)

FIG. 14

576 DDVMKKYEEFVCKWKQVGQPKLF. . . . . FATMDIEKCYDS..VNREK 615  
 : .:||| . .:| . .:| . .:| . .:| . .:| . .:| . .:| . .:|  
 379 NVLLKKVKH ANNLNLVSIPTQFNFDFYFVNQLQHLKLEFGLEPNILTKQK 426  
 516 LSTFL. .... KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKKEMK 657  
 | .  
 427 LENLLLSIKQSKNLKFLRLNFYTYYVAQETSRKQILKQATTIKNLKNNKNQ 476  
 558 DYFRQKFQKIALEGQQYPTLFSVLEN..EQNDLNAKKTLIVEAKQRNYFK 705  
 .  
 477 EETPETKDETPSESTSGMKFFDHLSELTELEDFSVN....LQATQEIY.. 520  
 706 KDNLLQPVINICQNYINFNGKFYKQTKGIPQGLCVSSILSSFYYATLEE 755  
 | .:  
 521 .DSLHKLLIRSTNLKKFKLSYKEMEKSMDTFIDLKNI.....YETLNN 564  
 756 SSLGFLRDESMNPENBNVNLLMRLTDDYLITTOENNANVLFIEKLINVSR 305  
 | .:  
 565 .  
 806 ENGFKFNMKKLQTSFPPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK 855  
 | .:  
 601 LQHAKYTFK..QNEFQFNNVKSAKIESSSLESLEDIDSCKSIASCCKNLQ 648  
 856 TLALMPNINLRLEGILCTLNLNMQT..KKASMLKK..KLKSFLMNNITH 901  
 .:  
 649 NVNI.....IASLLYPNNIQKNPFPNKPNNLFFFQFEQLKNLENVSINC 691  
 902 YFRKTI...TTEDFANKTLNKLFISSGGYKYMCAKEYKDHFKKNLAMSSM 948  
 .:  
 692 ILDQHILNSISEFLEKNKKIKAFILKRYYLLQYYLDYTKLFKTLQQLPEL 741  
 949 IDLEVSKIIYSVT.....RAFFKYLVCNIKDT..IFGEEHY 982  
 .:  
 742 NQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEFIKESSQTLQLIDF 791  
 983 PDFFLS TLKHFIGEISTKYY IFNRVCMLKAKEAKLKSQDCQSLIQ 1028  
 .:  
 792 DQNTVSDDSIKKILESISESKYHHYLRLNPSQSSSLIKSENEEIQELLK 840

*FIG. 14*  
(CONTINUED)

4 DIDLDDIENLLPNTFNKYSSSCSDKGCKTLKSGSKSPSLTIPK..... 47  
 : .:  
 617 NVKSAKIESSSLESLEDIDSCKSIASCCKNLQNVNIIASLLYPNNIQKNP 666  
 48 .....LQKOLEFYFSDANLYNDSFLRKLVLSGEQRVE....IETLLM 86  
 : .:  
 667 FNKPNNLFFFQFEQLKNLENVSINCILDQHILNSISEFLEKNKKIKAFIL 716

*FIG. 15*

17/59

1	MEMDIDLDIENL.....LPNTFNKYSSSCSDKKGCKTLKGSKSPS...	42
491	: ... .  ... .. : . .... .   ...: IELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKYGVRTCLEC	540
43	.LTIPKLQKQ.....LEFYFSDANLYNDSFLRKLVLKSGEQRVEIETLL  .: :   : :: ... : :: . .: :   ... :	85
541	ALVLGLMVKORCEKSSFYIFSSPSSOCNKCYL.EVDLPGDELRPSMOKLL	589

FIG. 16

telomerase p43 human La Xenopus LaA Drosophila La S. c. Lhp1p	LQKQLEF <del>Y</del> FSDANLYNDSFLRKVLVKSGEQRVEIETLLM ICHQEE <del>Y</del> FGDFNLPRDKFLKEQI.KLDEGWVWPLEIMIK ICEQIE <del>Y</del> FGDHNLPRDKFLKQQI.LLDDGWWVPLETMIK ILRQVE <del>Y</del> FGDANLNRDKF <del>Y</del> REQIGKNEGDGWMPLS <del>V</del> LVT CLKQVE <del>Y</del> FS <del>E</del> NFPYDRFLRTTAEK.NDGWMPISTIAT
---	---

FIG. 18

1	aactcattta	attactaatt	taatcaacaa	gattgataaa	aagcagtaaa	taaaaacccaa
61	tagatttaat	ttagaaagta	tcaattgaaa	aatggaaatt	aaaaacaact	aagcacaata
121	gccaaaagcc	aaaaaatgt	ggtggaaact	tgaatttagag	atgcaagaaa	acccaaatatga
181	tatataagg	agggttaaga	ttgacgatcc	taagcaatat	ctcgtgaacg	tcaactgcagc
241	atgttgg	taggaagtg	gttactacta	agataaaat	gaaagaagat	atatcatcac
301	taaagcactt	cttgagggt	ctgagtcgt	tcctgagttc	atctgttagt	tggcagtct
361	catccgtat	gaactttaca	tcagaactac	cactaactac	atgttagcat	tttggttgt
421	ccacaagaat	actcaaccat	tcatcgaaaa	gtacttcaac	aaagcagtac	ttttgcctaa
481	tgacttaactg	gaagtctgt	aatttgcata	ggttctctat	atttttgtat	caactgaatt
541	caaaaatttg	tatcttgata	ggatactttc	ataagatatt	cgtaaggaac	tcactttccg
601	taagtgttta	caaagatcg	tcagaagca	gttttctgaa	ttcaacaat	actaacttgg
661	taagtattgc	actgaatcc	aacgtaaagaa	acaatgttc	cgttacccct	cagttaacaa
721	caagtaaaaag	tgggattaaa	ctaagaagaa	gagaaaaagag	aatctttaa	ccaaaacttta
781	ggcaataaaaag	gaatctgaag	ataagtccaa	gagagaaaact	ggagacataa	tgaacgttga
841	agatgcata	aaggctttaa	aaccagcagt	tatgaagaaa	atagccaaga	gatagaatgc
901	catgaagaaa	cacatgaagg	cacccaaaaat	tccttaactct	accttggat	caaagtactt
961	gaccccttcaag	gatcttcatta	agtttgcctt	tatttctgaa	cttccaaaaaa	gagtcttaaa
1021	gatccttgg	aaaaaatacc	ctaagaccga	agaggaataac	aaagcagct	tttgttggatc
1081	tgcattctgc	cccttcaatc	ctgaatttggc	ttggaaagcgt	atgaagattt	aatctcttaa
1141	aacatgggaa	aatgaactca	gtgccaaaagg	caacactgt	gaggittggg	ataatttaat
1201	tccaagcaat	taactccat	atatggccat	gttacgtaa	ttgtctaa	tcttaaaaagc
1261	cgtgtttca	gatactacac	actttattgt	gtacaacaa	atttgcggc	ccaaaggccgt
1321	tgagaactcc	aagatgttc	tccttcattt	ctttagtgcc	atgaagctg	ttaatgaacg
1381	atttactaag	ggattcaagg	ccaaagaagag	agaaaaatatg	aatctttaaag	gtcaaatcg
1441	agcagtaaaag	gaagttgtt	aaaaaaaccga	tgaagagaaa	aaagatatgg	agttggagta
1501	aaccgaagaa	ggagaattt	ttaaagtcaa	cgaaggaatt	ggcaagcaat	acattaaact
1561	cattgaactt	gcaatcaaga	tagcgtttaa	caagaattta	gatgaaat	aaggacacac
1621	tgaatcttc	tctgtatgtt	ctgggtctat	gagtacttca	atgtcagg	gagccaagaa
1681	gtatggttcc	gttctgtact	gtctcgagt	tgcattatgc	cttgggttga	tggaaaata
1741	acgttgtgaa	aagtccctat	tctacatctt	cagttcacct	agttctcaat	gcaataatgt
1801	ttacttagaa	gttgatctcc	ctggagacga	actccgtct	tctatgtaaa	aacttttgca
1861	agagaaaagg	aaacttgg	gtggtaactga	tttccccctat	gagtgcatt	atgaatggac
1921	aaagaataaaa	actcacgt	acaatatctgt	tattttgtct	gatatgtat	ttgcagaaagg
1981	attatccat	atcaatgtt	gaggcagt	cattttgttac	agcataaaaa	agtacaagg
2041	tgaagtaaaat	cctaaacatta	aaatcttgc	agttgactt	gaaggttacg	gaaagtgcct
2101	taatcttagt	gatgagttca	atgaaaacaa	ctacatcaag	atattcggt	tgagcgttcc
2161	aatcttaaaag	ttcatttcag	ccaaagcaagg	aggagcaat	atggtcgaag	ttatcaaaaa
2221	cttgcctt	aaaaaaatag	gacaaaatgt	agtttcttgc	gattcttcta	taacaaaaat
2281	ctcacccccc	ttttttttt	tattgtatag	ccattatgaa	atthaattt	ttatcttatt
2341	atttaagtt	cttacatagt	ttatgtatcg	cagtcatttca	gcctatttca	atgattctgc
2401	aaagaacaaa	aaagattaaa	a			

*FIG. 19*

+

18/59

	Motif A	Motif B	Motif C	Motif D	Motif E
Consensus	h--hdh---h--h	h---+---OG---SP	h-YnDnhh	Gh-h---K	h-hGh-h
telomerase p123	GQPKLFFATMDIEKYCDSTAREKLSTTELKTTKLL-100-KFYKQTKGIPQGLCVSSILSSPFYATLEESSLGFL	KNRNLHACTDYDKRAFDSTPHSMWLIQVLEIYKIN-	-14-LMRLTDYDYLITTTQENN-O-AVLFIEKLINVSRENGFKFMKKLQT-23-QDYCDWIGISI		
Dong (LINE)	KNRNLHACTDYDKRAFDSTPHSMWLIQVLEIYKIN-	28-RQIAIKKGTYQENDS-SPIWMCLAINPLSHQLHNDR	-16-HLYMDIDIKLYARNKE-O-MKKLIDDTTIFSNLDISMQFLGLDKCKT-25-KLYKTYLGFQQ		
a1 S. c. (groupII)	FGGSNWMFREVDLKRKFDTLSDHLLIKELKRYISD-	26-HVPVGPRVCYQGAPTSIPALCNAVLLRLDRRLAGLA	-55-YVRYADDILIGVLGSKN-2-KLTKRDLNINFNS .LGLTINEEKILLI- 4-ETPARELGYN		
HIV-RT	LKKKKSVTIVLDDYDATESYVPLDEDKRYTAFTIP-	27-GIRYQYNVLPQGMKGS-SATFQSSMTKILEPFRKON	- 4- TYQYMDIYVGSHIETG-1-HRTKIEELRQHLLRMGLTPDKKHQK- 0-EPPFLMGYEL		
L8543.12	VLPELYFMKFDVKSCTDS1PRMECMRLKDALKN-	68-KCYIREDGLFQGSSSLAPIVLDLLEYSEFK	- 8-ILKLADDFLITSTDQQQ-.....VINTIKKLAMGGFOKYNAKANR-41-IRSKSSKGFR		

FIG. 17

19/59

MEIENNQAQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL  
VNVTACLLQEGSYQDKDERRYIITKALLEVAESDPEFICQLA  
VYIRNELYIRTTTNYIVAFCVVHKNTQPFIEKYFNKAVLLPNDL  
LEVCEFAQVLIFYDATEFKNLYLDRILSQRDIRKELTFRKCLQRC  
VRSKFSEFNEYQLGKYCTESQRKKTMFRLSVTNKQKWDQTKKK  
RKENLLTKLQAIKESEDKSRETGDIMNVEDAIKALKPAVMKKI  
AKRQNAMKKHMKAQKTPNSTLESKYLTFKDLIKFCHESEPKERV  
YKILGKKYPTEEEYKAAGDSASAPFNPELAGKRMKIEISKTW  
ENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGVSDTT  
HSIVINKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKAKKR  
ENMNLLKGQIEAVKEVVEKTDEEKKDMELEQTEEGEFVKVNEGIG  
KQYINSIELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGA  
KKYGSVRTCLECALVGLMVQRCERKSSFYIFSSPSSQCNKCYL  
EVDLPGDELRPSMQKLLQEKGKLGGSDFPYECIDEWTKNKTHV  
DNIVILSDMMIAEGYSDINVRGSSIIVNSIKKYKDEVNPNIKIFA  
VDLEGYGKCLNLGDEFNENNYIKIFGMSDSILKFISAKQGGANM  
VEVIKNFALQKIGQK

FIG. 20

0  
9  
8  
7  
6  
5  
4  
3  
2  
1

MSRRNQKKPQAPIGNETNLDVLQNLEVYKSQIEHYKTQQQQIK  
EEDLKLLKFKNQDQDGNGSNDDDEENNSNKQQELLRRVNOIQO  
QVOLIKKVGSKVEKDLNLNEDENKNGLSEQVKEEQLRTITEE  
QVKYQNLVFNMHDYQLDLNESGGGHRRRRETDYDTEKWFEISHDQ  
KNYVSTYANQKTSYCWMLKDYFNKNNDHHLVSINRLETEAEFY  
AFDDFSQTICKLTNNSYQTVNIDVNFDNNLCTILLRFLLSLERF  
NILNIRSSYTRNQNFKEKIGELLETIFAVVFSHRHLQGIHLQVP  
CEAFQYLVNSSSQISVKDSQLOVYSFSTDLKLVDTNKVQDYFKF  
LQEFPRLTHVSQQAIPVSATTAVENLNVLKKVKHANLNLSIP  
TQFNFDYFVNQLQHLKLEFGLEPNILTQKLENLLSIQSKNL  
KFLRLNFYTYVAQETSRKQILQATTIKNLKNNKNQEEETPETKD  
ETPSESTGMKFFDHLSELTELEDFSVNLQATQEIYDSLHKLLI  
RSTNLKKFKLSSYKYEIMEKSKMDTFIDLKNIYETLNNLKRCGVNI  
SNPHGNISYELTMKDSTFYKFKLTLNQELQHAKYTFKQNEFQFN  
NVKSAKIESSSLSESLEDIDSCKSIASCKNLQNVNIIASLLYPN  
NIQKMPFNKPNLFFKQFEQLKNLENVSINCILDQHILNSISEF  
LEKNKKIKAFLKRYLLQYLDYTKLFKTQQLPELNQVYINQ  
QLEELTVSEVHKQVWENHKQAFYEPLCEFIKESSQTLQLIDFD  
QNTVSDDSIKKILESESKEYHHYLRLNPSQSSSLIKSENEEIQ  
ELLKACDEKGVLVKAYYKFPCLPTGTYYDYNNSDRW

FIG. 22

MKILFEFIQDKLIDLQTNSTYKENLKGHFNGLDEILTTCFAL  
PNSRKIALPCLPGDLHSKAVIDHCIIYLLTGELYNNVLTFGYKI  
ARNEDVNNSLFCHSANVNVTLLKGAAWKMFHSLVGTYAFVDLLI  
NYTVIQFNGQFFTQIVGNRCNEPHLPPKVVQRSSSSATAAQIK  
QLTEPVTKQFLHKLNNINSSFFPYSKILPSSSSIKKLTDLREA  
IFPTNLVKIPQRLKVRINLTQKLLKRHKLNRVYVSILNSICPPL  
EGTVLDSLHSRQSPKERVLKFIIIVILQKLLPQEMFGSKKNKGK  
IIKNLNLLSLFNLNGYL\_PFDSSLKKLRLKDFRWLFISDIWFTKH  
NFEFLNQLAICFISWLFRQIIPKIIQTFYCTEISSTVTIVYFR  
HDTWNKLITPFIIVEYFKTYLVEENNVCRNHSYTLSNFNHSKMR  
IPKKSNNEFRIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEY  
LRNKRPTSFTKIQSPTQIADRIKEFKQRLKKFNNVLPPELYFMK  
FDVKSCYDSIPRMECMRILKDAKLNENGFFVRSQYFFNTNTGVL  
KLFNNVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTA  
LWVEDKCYIREDGLFQGSSLASAPIVDLVYDDLLEFYSEFKASPS  
QDTLILKLADEFLIISTDQQQVINIKKLAMGGFQKYNNAKANRDK  
ILAVSSQSDDDTVIQFCAMHIFVKELEVWKSSTMNFHIRSKS  
SKGIFRSLIALFNTRISYKTIIDTNLNSTNTVLMQIDHVVKNISE  
CYKSAFKDLSINVTONMQFHSFLQRIIEMTVSGCPITKCDPLIE  
YEVRFTILNGFLESLSSNTSKFKDNIILLRKEIQHLQAYIYIYI  
HIVN

FIG. 23

+

20/59

1	tcaatactat	taattaataa	ataaaaaaaa	gcaaaactaca	aagaaaatgt	caaggcgtaa
61	ctaaaaaaag	ccataggctc	ctataggcaa	tgaaacaaat	cttgattttg	tattacaaaa
121	tctagaagtt	tacaaaagcc	agattgagca	ttataagacc	tagtagtaat	agatcaaaga
181	ggaggactcto	aagcttttaa	agtccaaaaa	ttaaaggat	gatgaaact	ctggcaacgc
241	tgtatgtat	gaagaaaaca	actaaaaata	ataataagaa	ttataagga	gagtcaattt
301	gattaatgt	caagttaat	tgataaaaa	agttggttct	aaggtagaga	agatgttcaa
361	tttgaacgaa	gatgaaaaca	aaaagaatgg	actttctgaa	tagcaagtga	aagaagagta
421	attaagaacg	attactgaag	aataggttaa	gtattaaaat	tttagtattt	acatggacta
481	ccagttatag	ttaaatgaga	gtggggcca	tagaaagacac	agaagagaaa	cagattatga
541	tactggaaaa	tgggtggaaa	tatccatgaa	caaaaaaaaat	tatgtatcaa	tttacgcca
601	cttaaaaagaca	tcatatgtt	gggtggcttaa	agattatttt	aataaaaaca	attatgatca
661	tcttaatgt	agcattaaca	gactagaaaac	tgaagccgaa	ttctatgcct	ttgatgattt
721	ttcacaaaca	atcaaactta	ctaataattc	ttactagact	gtaaacatag	acgttaattt
781	tgataataat	ctctgtat	tcgcattgt	tagattttt	tttactactag	aaagattcaa
841	tatTTGAT	ataagatctt	cttataacaag	aaataatatt	aattttgaga	aaattgttga
901	gctacttggaa	actatctcg	cagggtgtt	ttctcatcgc	cacttacaag	gcattcattt
961	acaagttcct	tgcgaagcgt	tctaataattt	agttactcc	tcatcataaa	tttagcgtt
1021	agatagctaa	ttataggat	actctttctc	tacagactta	aaatttagtt	acactaaca
1081	agtccaaagat	tatTTTAA	tcttataaga	attcccctcg	ttgactctat	taagcttagt
1141	ggttatcccc	gttagtgc	ctaagcgtgt	agagaacctc	aatTTTTAC	ttaaaaagggt
1201	caagcatgct	aatcttaatt	tagttctat	ccctacttac	ttcaatttttgc	atttctactt
1261	tgttaattt	taacattga	aatttagagtt	tggtttagaa	ccaaatattt	tgacaaaaca
1321	aaagcttgaa	aatctacttt	tgagtataaa	ataatcaaaa	aatcttaat	tttaagattt
1381	aaacttttac	acctacgtt	cttaagaaaac	ctccagaaaa	cagatattaa	acaagctac
1441	aaacaatcaa	aatctcaaaa	acaataaaaa	tcaagaagaaa	actctgttt	ctaaagatgt
1501	aactccaaag	gaaagcaca	gtgtttagaa	atTTTTGAT	catTTTTCTG	atttaacccg
1561	gcttgaagat	ttcagcggtt	acttgtaagc	taccccaagaa	attttatgata	gcttgcacaa
1621	acttttgatt	agatcaacaa	atttttttttt	gttcaaatta	agttacaaat	atgaaatgg
1681	aaagagttaa	atggatacat	tcatacatct	taagaatatt	tatggaaacct	taaacaatct
1741	taaaagatgc	tctgttata	tatccaaatcc	tcatggaaac	atTTTTCTG	actgtacaaa
1801	taaaagattt	actttttt	aattttat	gaccttaaac	tttttttttt	aacacgcttac
1861	gtatactttt	aatTTTTA	aattttat	taataacgtt	aaaagtgc	aaatttgatc
1921	ttcctcatta	gaaagcttag	agatattga	tagtcttgc	aaatcttattt	tttcttgc
1981	aaatttacaa	aatgttata	ttatcgccag	tttgccttat	cccaacaata	tttagaaaaaa
2041	tccttcaat	aaggccaaatc	ttcttattttt	caagcaattt	gaataatttga	aaaatttg
2101	aaatgtatct	atcaactgt	ttcttgcata	gcatataactt	aattttat	cagaattttt
2161	agaaaaaaat	aaaaaaaataa	aaggccat	tttggaaaaga	tattttat	tacaatattt
2221	tcttgattat	actaaattat	ttaaaacact	tcaatagttt	ccttgcattaa	attaagttt
2281	cattaattag	caatttgc	aatttgcattt	gagtgttgc	cataatgtt	tatggggaaa
2341	ccacaagcaa	aaagctttct	atgaaaccatt	atgtgttttt	atcaaaaat	cattccaaac
2401	cccttgcata	atagatttt	acccaaaacac	tgttgcattt	gacttcttta	aaaagatgt
2461	agaatctata	tctgttgc	atgtatcatca	tttttttttt	tttttttttt	cagaattttt
2521	cagtttattt	aaatctgaaa	acgaaagaaaat	ttaaagaactt	ctcaagctt	gacgacgaaa
2581	agttgtttt	gttttttttt	actataattt	ccctctatgt	ttaccaactg	gtacttattt
2641	cgatttacat	tcagatagat	gggttatttt	taataatattt	tttttttttt	tattttat
2701	tgtatattt	tttgcattt	attttttttt	tacatataat	agtctttttt	agtgttttga
2761	atatttttta	gttattttat	tcattttttt	aagttaaataa	ttttttttca	atcatttttca
2821	aaaaaaatcg					

FIG. 21

21/59

Oxytricha  
Euplotes

LCVSYILSSFYANLEENALQFLRKEAMDPEKPETNLLMRLT  
LCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLT

**FIG. 24**

ATTTATACTCATGAAAATCTTATTGAGTTCAAGACAAGCTTGACATTGATCTACA  
GACCAACAGTACTTACAAGAAAATTAAATGTGGTCACTTCAATGGCCTCGATGAAAT  
TCTAACTACGTGTTGCACCAATTCAAGAAAATAGCATTACCATGCCCTCCCTGG  
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCTTACTGTTGACGGGCGAATT  
ATACAACAAACGTAACATTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAAATAG  
TCTTTTTGCCATTGCAAATGTTAACGTTACGTTACTGAAAGGCCTGCTTGAAAAT  
GTTCCACAGTTGGCTGACATACGCATTGTTATTGATCAATTATACAGTAAT  
TCAATTAAATGGGCAGTTTCACTCAAATCGTGGGTAAACAGATGTAACGAACCTCATCT  
GCCGCCAAATGGGTCCAACGATCATCCTCATCATCGCAACTGCTGCGCAAATCAAACA  
ACTTACAGAACCAAGTGACAATAAACATTCTTACACAAGCTCAATATAAATTCCCTTC  
TTTTTTCTTATAGCAAGATCCTCCTTCATCATCATCTATCAAAAGCTAACTGACTT  
GAGAGAAGCTATTTCACAAATTGGTAAATTCCTCAGAGACTAAAGGTACGAAT  
TAATTGACGCTGCAAAGCTATTAAAGAGACATAAGCGTTGAAATTACGTTCTATT  
GAATAGTATTGCCCCACCATTGGAAAGGGACCGTATTGGACTTGTGCGATTGAGTAGGCA  
ATCACCAAAGGAACGAGCTTGAAATTATCATTGTTATTACAGAAGTTATTACCCCA  
AGAAATGTTGGCTCAAAGAAAATAAGGAAAATTATCAAGAATCTAAATCTTTATT  
AAGTTTACCCCTTAAATGGCTATTACCATTTGATAGTTGTTGAAAAAGTAAAGATTA  
GGATTTCGGTGGTTGTTCAATTCTGATAATTGGTICACCAAGCACAATTGAAAACCTT  
GAATCAATTGGCGATTGTTCAATTCCCTGGCTATTAGACAACATAATTCCAAAATAT  
ACAGACTTTTACTGCACCGAAATATCTTACAGTGACAATTGTTACTTAGACA  
TGATACTGGAATAAACTTATCACCCCTTTATCGTAGAATATTAAAGACGTACTTAGT  
CGAAAACACGTATGAGAACCCATAATAGTTACACGTTGTCATTCATAGCAA  
AATGAGGATTATACCAAAAAAAAGTAAATGAGTTCAAGGATTATTGCCATCCCAG  
AGGGCAGACGAGAAGAATTACAATTATAAGGAGAATCACAAAATGCTATCCAGCC  
CACTCAAAATTTAGAATACCTAAGAAACAAAGGCCGACTGTTACTAAATATA  
TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTAAAGCAGAGACTTTAAAGAAATT  
TAATAATGTCTTACAGAGCTTATTCAATGAAATTGATGTCAAATCTGCTATGATT  
CATACCAAGGATGGAATGTATGAGGAACTCAAGGATGCGCTAAAAATGAAATGGGTT  
TTCGTTAGATCTCAATATTCTCAATACCAATACAGGTGTATTGAAGTTATTAAATGT  
TGTTAACGCTAGCAGAGTACCAAAACCTTATGAGCTACATAGATAATGTGAGGACGGT  
TCATTATCAAATCAGGATGTTAAACGTTGAGAGATGAAATATTAAACAGCTTT  
GTGGGTTGAAGATAAGTGCACATTAGAGAAGATGGCTTTTCAGGGCTCTAGTTATC  
TGCTCCGATCGTGTGGTATGACGATCTCTGGAGTTTATAGCGAGTTAAAGC  
CAGTCTCAGGCCAGGACACATTAATTAAACTGGCTGACGATTCTTATAATATCAAC  
AGACCAACGCAAGTGAATATCAAAAGCTTGCATGGCGGATTCAAATATAA  
TGCAGAAAGCCAATAGAGACAAAATTAGCCGTAAGCTCCCAATCAGATGATGATACGGT  
TATTCAATTGTCGAATGCACATATTGTTAAAGAATTGAAAGTTGAAACATTCAACAG  
CACAATGAATAATTCCATATCCGTTGAAATCTAGTAAAGGGATATTGCAAGTTAAT  
AGCGCTTTAACACTAGAATCTTATAAAACAATTGACACAAATTAAATTCAACAA  
CACCCTCTCATGCAAATTGATCATGTTGAAAGAACATTGCGAATGTTATAATCTGC  
TTTTAAGGATCTATCAATTATGTTACGCAAATATGCAATTTCATTGTTCTTACAACG  
CATCATTGAAATGACAGTCAGCGGTTGTCACATTACGAAATGATCCTTAAATCGAGTA  
TGAGGTACGATTCCACCATATTGAAATGGATTGGAAAGCCTATCTTCAAACACATCAA  
ATTAAAGATAATTCAATTCTTGTGAGAAAGGAAATTCAACACTGCAAGC

**FIG. 26**

human	ISEIIEWLVLGKRSNAKMCILSDFEKKOQIAFAETYWLNYSFLIPILQSFYVTEITSSDLRNR	AKFLHWLMSYYVELRSFFYVTEITTEQKMR
tez1	LKDFRWLFIISD--JWFTKHNFENLNOLAICFISWLFFQLIKPQITQFFYCTEISSTV-	Motif 1
EST2	TREI SWM{QVET-SAKHFYTFDHEN-IYVLMKLLRWFEDUVSLIRCFYVTEQKSYSK	*
p123	*	*
human	LFYFYRKSWWSKLQSQSTIGIROQHLKRVQLRDVSSEAEVROHEARPALLTSRLRFIPKP--DGL	
tez1	TVYFRKDIWKLLCRPFI-TSMKMEAFEKINENNVRMDIQK-TTLPAVTRLLPKK--NTF	
EST2	IVYFRHDITWNKLITTPFIVEFKTYLVENNVCRNENSYTLLS--NFNHSKMRILIPKKSNNEF	
p123	YYYYRKNTIWDMKMSI-ADLKETLAEVQEKEVEEMKKS-LGFAPGKLRLLIPKK-TTF	*
human	RPTVMDDYVVGARTFRREKRAERLTSRVKALF-SVLYNTERA	Motif 2
tez1	RLITN-LRKRFELIKMGNSKKMLVSTNOTLRPVAVASILKHLNEESESGIPFNLEVYMKLLITF	
EST2	RJIALPCRGADEEEFTTYKENHKNAIQOPTOKILEYRNRKRTPSFTKIKYSPTQIADRIKEF	
p123	RPTMTFNKKIVNSDRKTTKLTTNTKLNNSHMLKTLYLN-RMFKDPFGFAVFNYYDDVMKCY	*
human	KKDLLKHMGR-KKYFVRIDIKSCYDRIKQDLMFRVKK-KLKDPFVFIRKYATIHATS	Motif 3 (A)
tez1	KORLLKKFNNVLPPELYFMKFDVKSCYDSIFPRMCEMRLKD-ALKNENGFFFVRSQYQFFNTN	
EST2	EFFVCKWVKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLSSDFWIMTAQILKRN	
p123	*	*

FIG. 25

23/59

AKFLHWLMSVYVVELRSFFYVTETTFQKNRLFFYRKSVWSKLQSIGIRQHLKR  
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMMDYVVGARTFRREK  
RAERLTSRVKALFSVLNYERA

FIG. 27

GCAGTTCCCTGCACTGGCTGATGAGTGTACGTCGTCGAGCTGCTCAGGTC  
TTCTTTATGTACGGAGACCACGTTCAAAAGAACAGGCTCTTTCTACCG  
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACCTGAAG  
AGGGTGCAGCTGCCGGACGTCGGAAAGCAGAGGTCAAGGCAGCATCGGGAAAGC  
CAGGCCCCCCCCTGCTGACGTCAGACTCCGCTTCACTCCCAAGCCTGACGGGC  
TGCGGCCGATGTGAACATGGACTACGTCGTGGAGCCAGAACGTTCCCGAGA  
GAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCAGTGTTCAGCGTGCT  
CAACTACGAGCGGGCGCG

FIG. 28

MTEHHTPKSRILRFLLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSVDVQTFSISFLHSTVVGFD  
DSKPDEGVQFSSPKCSQSELIANVVQMFDESERRNLLMKGFSMNHEDFRAMHVNQNDLVSTF  
PNYLISILESKNWQLLEIIGSDAMHYLLSKGSIFEALPNDNYLQIISGIPLFKNNVFEETVSKKRKR  
TIETSITQNKSRKEVSWNSISISRFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG  
LINAQVKQLHKVIPLVSQSTVVPKRLLKVPILLEQTAKRHLRISLSKVYNHYCPYIDTHDDEKILS  
YSLKPNQVFAFLRSILVRVPKLIWGNQRIFEIIILKDETLFLKLSRYESFLHYLMSNIKISEIEWL  
VLGKRSNAKMCLSDFEKRQIFAEFIYWLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKL  
PFITSMKMEAFAEKINENNVRMDTQKTTLPAPAVIRLLPKKNTFRLITMLRKRFLIKMGSNKKMLVSTN  
QTLRPVASILKHLNEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDI  
KSCYDRIKQDLM  
FRIVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVQVQLLSMKTSDTLFVDFV  
WTKSSSEIFFKMLKEHLSGHIVKIGNSQYLQKVGIPQGSILSSFLCHFYMEDLIDEYLSFTKKKG  
SVL  
LRVVDDFLIFTVNKKDAKKFLNLSLRGFEKHNFSTSLEKTVINFENSNGIINNTFFNESKKRMPFFG  
FSVNMRSLDTLLACPIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCN  
IYRLGYSMCMRAQAYLKRMDIFIPQRMFITDLLNVIGRKIWKKLAEILGYTSRRFLSSAEVKWLFC  
LGMRDGLKPSFKYHPCFEQSLTDLIKPLRPVLRQVLFLHRIAD

FIG. 29

On the 1<sup>st</sup> of January, 1863, the following resolutions were adopted by the General Assembly of the Commonwealth of Massachusetts:

+

24/59

FIG. 30

*FIG. 30*  
(CONTINUED)

+  
EST2 pep  
Euplotes pep  
Trans of tetrahymen  
Consensus

FFYCTEISST VTIIVFRHDT WN---KLIT P---FIVE YFK-TYLVEN  
FFYVTEQQKS YSKVYYRKN IWDVI-MKMS TAD---LKK ETLA--EVQE  
-----KHKE GSQIIFYYRKPK WKRVLVSKLTIVKVRIFQSEK NKQMKNNFYQ  
40  
43  
44  
FFY.TE.. K. . S .. YYRK. IW... -KL... . . . . F.. K . . . . V...  
50

NVCRNHNSY- ----- TLSNPNHSK RIIIPKSNNE FRI[IAIPCRG  
KEVEENKSL ----- GFAPKGS RIIPKKT- FRPIIMTFNKK  
KIQLEENLE KYEEKLIPED SFQKYPQGKL RIIPKGS-- FRPIIMTFLRK  
K...E..... ----- . . . . F.. GKL RIIPK. . -- FRPIIMTF.RK  
79  
78  
92  
100

ADEEEFTIYK ENHKNAIQT QKILEYTRNK RPTSFTKIYS PTQIADRIKE  
IVNSDRKTTK LTNTNKLNS HILMLKTLKN- -----RMFK -DPFGFAVEN  
DKQKNIK --- LNLNQNLMDS QLVERNLKD- -----ML-G -QKIGYSVFD  
.....K..K LN.N..L..S QI.L..LKN- ----- . . . . . . . .  
129  
120  
130  
150

FKQRLLKKFN NVL----- -PEIYFMKFD VKSCYD  
YD-DVKMKVIE EFVCKMKQVH CPKLFATMD IEKYCD  
NK-QISEKFA QFIEKMKNG RPCLQYYVTL- -----  
.K-...KKF. .F..KWK..G .P..LYF.T.D . . .CYD  
157  
155  
158  
186

+  
FIG. 31

27/59

S-1: FFY VTE TTF QKN RLF FYR KSV WSK  
S-2: RQH LKR VQL RDV SEA EVR QHR EA  
S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q  
A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS  
A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

*FIG. 32*

OPEN SOURCE MATERIAL

Poly 4  
t t t c  
t a a g c c t c g  
5' - c a g a c c a a a g g a a t t c c a t a a g g - 3'  
Q T K G I P Q G

4(B')

5(c')

D D Y L L I T  
3' - c t g c t g a t g g a g g a g a t a g t g g - 5'  
a a a a a a a a  
t t t t  
c c  
Poly 1

*FIG. 34*

+

+

28/59

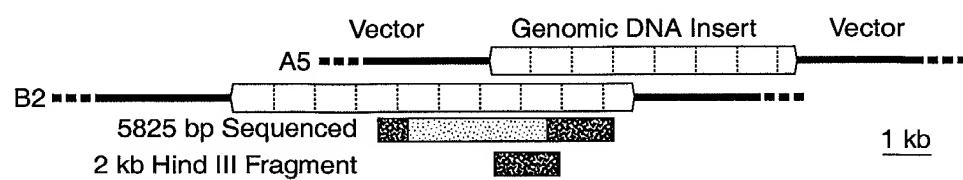


FIG. 33A

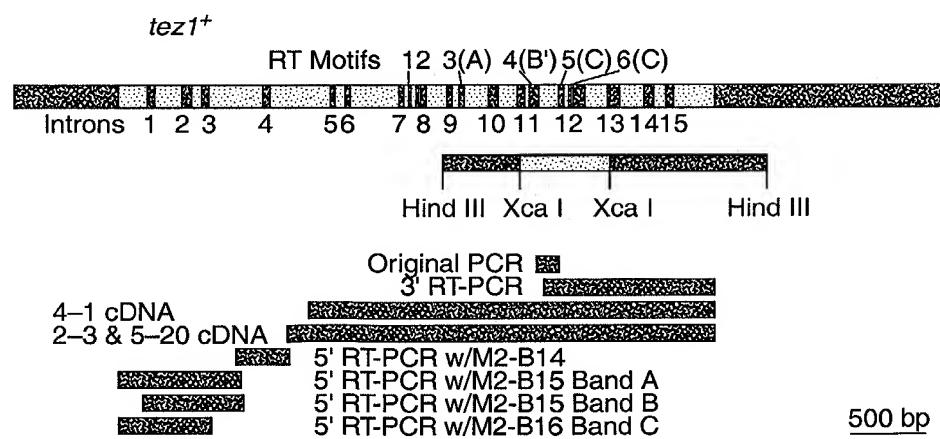
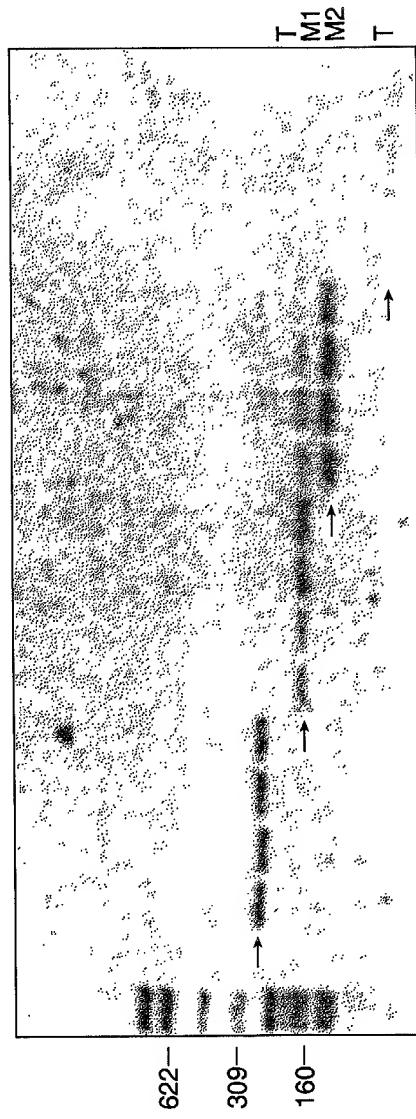


FIG. 33B

+

+

29/59



Motif C (5)  
DDYLLT

Motif B' (4)  
QTKGIPQQ

FIG. 35

+

30/59

FIG. 36

+

31/59

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA  
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT  
E D L I D E Y L S F T K K G S V L L R

GTA GTC gac tac ctc ctc atc acc  
CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tgg  
a a a a a a a a

t t t t t t t t

c c c c

Poly 1

.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence  
D D F L F I T

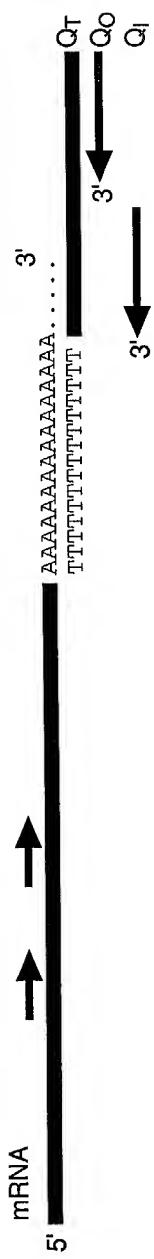
FIG. 36  
(CONTINUED)

+

...TTTTTTTTTTTTTTT

+

32/59



1. Synthesis of cDNA with  $Q_T$  Primer



2. First Round PCR Using Outside Primer and  $Q_O$  Primer



3. Second Round PCR Using Inside Primer and  $Q_I$  Primer



4. Sequence Second Round PCR Products Using Inside Primer  $Q_I$  Primer

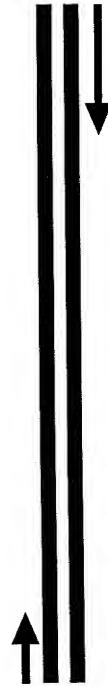


FIG. 37

+

+

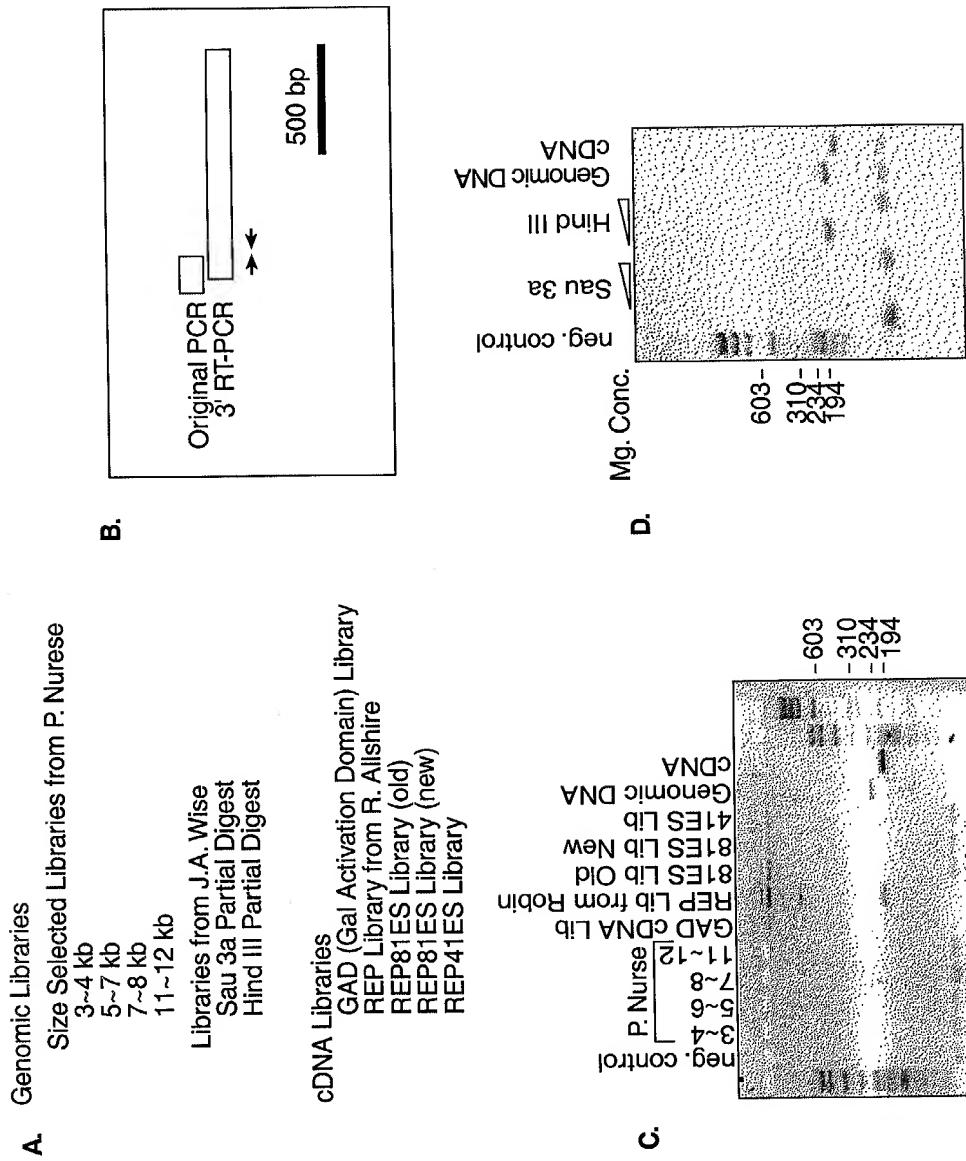


FIG. 38

+

0976626 - DEPDI

+

34/59

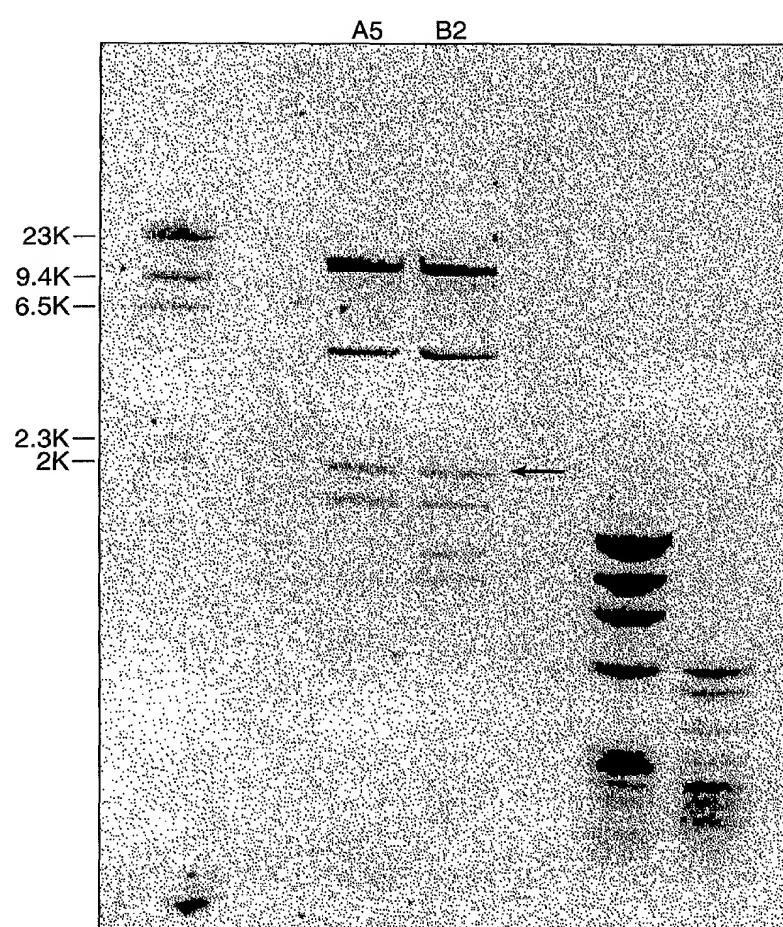


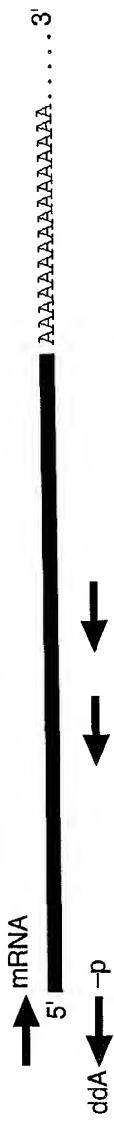
FIG. 39

+

Y. C. D. S. C. S. S. S. S. S. S. S. S. S.

+

35/59



1. Synthesis of cDNA with Specific Destination Primer



2. Ligate Oligo with 5' -P and blocked 3' to cDNA using T4 RNA Ligase



3. First Round PCR



4. Second Round PCR



FIG. 40

+

+

+

+

			Motif 0
S.p.	Tez1p	(429)	WLYNSFLIPILOQFFYITESSDLRNRTVYFRKDIW ... (35) ...
S.c.	Est2p	(366)	WLFRQLIPKIIQTFFYCTEBISSTVT-IVYFRHDTW ... (35) ...
E.a.	p123	(441)	WIFEDLVVSLIRCFYYTQQKSYSKTYYRKNIW ... (35) ...
	*	***	*
			Motif 1
			Motif 2
			K
			R
S.p.	Tez1p	AVIRLLPKK--NTFRLITN-LRKRF	... (61) ...
S.c.	Est2p	SKMRLIIPKKSNNNEFRILIAIPCRGAD	... (62) ...
E.a.	p123	GKLRLIPKK--TTFRPIMTFNKKIV	... (61) ...
	*	***	*
			Motif 3 (A) AF
S.p.	Tez1p	KKYFVRIDLKSCYDRIKODLMFRIVK	... (89) ...
S.c.	Est2p	ELYFMKFDVKSCYDSSIPRMECMRILK	... (75) ...
E.a.	p123	KLFATMDIEKCYDSVNREKLSTFLK	... (107) ...
	*	***	*
			Motif 4 (B')
			hPQG pp hh h
S.p.	Tez1p	YLQKVGTPOQGSILSSFLCHFYMEDLIDEYLSF	... (6) ...
S.c.	Est2p	YIREDGFLFOQGSSLsapIVDLYVDDLEFYSEF	... (8) ...
E.a.	p123	YKQTKGIPQGLCVSSILSSEFYATLEESSLGF	... (14) ...
	*	**	*
			Motif 5 (C)
			h F DDhh
S.p.	Tez1p	VLLRVVDDFLFITVNKKDAKKFLNLSLRGFEKHNFSSTSLEKTVINFENS	... (205) ...
S.c.	Est2p	LILKLAADDFLIISTDQQQVINTIKLAMGGFQKYNAKANRDKILAVSSQS	... (173) ...
E.a.	p123	LIMRLTDYLLITTOQENNNAVLFIEKLINVSRENGFKENMKKLQTSFPLS	... (209) ...
	*	**	*

36/59

FIG. 41

A.	Sp_Tip1p	1	- - - - -	M T E H H T P K S R I L R F L E N Q Y V Y L C T	24
	Sc_Est2p	1	- - - - -	M K   L F E F	7
	Ea_p123	1	M E V D V D N Q A D N H G I H S A L K T C E E   K E A K T L Y S W	33	
Sp_Tip1p	25	L N   D Y V Q L V   L R G S P A [ S Y S N I C E R   L R S D V Q T S F S	57		
Sc_Est2p	8	I Q D K L D   D L Q T N - - S T Y K - - E N   L K C G H F N G L D	35		
Ea_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K   F A Q T N	61		
Sp_Tip1p	58	I F   L H S T V V G F D S K P D E G V Q Q F S S P K C S Q S E L   A N	90		
Sc_Est2p	36	E I   L T C F A L P N S R - K I A L P C L P G D L S H K A V I D H	67		
Ea_p123	62	I V A T P R D Y N E E D F K V   A R K E V F S T G L M I E L D K	94		
Sp_Tip1p	91	V V K Q M F D E S F E R R   R - N   L M K G F S M N H E D F R A M H	122		
Sc_Est2p	68	C I   Y   L L T G E L Y N - - N   V   L T F G Y K I A R N E D - - -	93		
Ea_p123	95	C L V E L L S S S D V S D R Q K   L Q C F G F Q L K G N Q - - -	122		
Sp_Tip1p	123	V N G V Q   N D   L V S T F P N Y L I S !   L E S K N W Q L L L E I   G	153		
Sc_Est2p	94	- - V N N   S   L F C H S A N V N V T   L K G A A W K M F H S L V G	122		
Ea_p123	123	- - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M   G	153		
Sp_Tip1p	156	S D   A M H Y   L S K G S   F E A L P N D N Y L Q   S G   P   F E K N	188		
Sc_Est2p	124	T Y A   F V D   L L I N Y T V I Q F N - G Q F F T Q   I V G N R C N E P	156		
Ea_p123	153	N E   F R H   L Y T K Y L I F Q R T S E G T L V Q   F C G N N V F D H	188		
Sp_Tip1p	189	N V F E E T V S K K R K R T   E T S   T Q N - - - K S A R K E V S	219		
Sc_Est2p	156	H L P P K W V Q - - R S S S S A T A A Q I - - - K Q L T E P V T	188		
Ea_p123	186	L K V N D K F D K - - K Q K G G A A D M N E P R C C S T C K Y N V K	211		

FIG. 42

A.

Sp_Tip1p	219	WNSISISRFSSIFYRSSYKKFKQDLYFNLHSICD	251
Sc_Est2p	184	N-----NEK-----DHFLNNINVPNWNNMKSRTRIFYCTHEN	200
Ea_p123	218	-----N-----R-----	248
Sp_Tip1p	252	RNTVHMMWLQWIFPRQFGLINAFQVKQLHKVIPL	284
Sc_Est2p	201	-----YSKILPSSS-----SIKKLTDLREAIFP	223
Ea_p123	249	R-----NNQFFIKKHEFVSNKNNISAMDRAQTI	275
Sp_Tip1p	285	VST-----QS-----VVPKRLLKLYPLIEQTAKRLLHRIS	313
Sc_Est2p	224	TN-----LVKIPQRBLKVRI-----TLQKLLKRHKRNLN	252
Ea_p123	276	FTNIFRFRNIRKQLKDVKYEKIAYMLLEKVKDFN	308
Sp_Tip1p	314	LSKVVYNHYCPTYID-THDDEKILSYSLKPNQ-----	342
Sc_Est2p	253	YVSILNSICPPLEFGITVLSDLSHLSRQSPKER-----	282
Ea_p123	309	FNYYLTKSCPLPENWRERKQKINENLINKTREEK	341
Sp_Tip1p	343	-----VFAFLRSILVVRVFPKLI-----	359
Sc_Est2p	283	-----VLIKFI-----VILQKLLPQEM-----	299
Ea_p123	342	SKYYEELFSYTDDNKCYTQFINEFFYNILPKDF	374
Sp_Tip1p	360	WGNGQRIFEEILKDLLETFLKL-----	392
Sc_Est2p	300	FGSKKNKGKIKNLNL-----PLNGYLPFDSSLK	332
Ea_p123	375	LTG-RNRKNFQKKVKKYVELNKHIELIHKNLLLE	406
Sp_Tip1p	393	NIKISEIEWLVLGKRSNAKMCLSDFEKRKQIFA	425
Sc_Est2p	333	KLRLKDFRWLFIS----DIWFTEKHNFENLNQLA	362
Ea_p123	407	KINTREISWMQVETS-AKHFYYFDHEN-IYVLW	437

A.		
Sp_Tip1p	426	EF YWLYNSF I PILQSFFY ITESSDLRNRTVY  458
Sc_Est2p	363	CF SWLFRQL PKI QTFFYCTESTVT-IYY  394
Ea_p123	438	KLLRWIEEDLVVSLIRCFFY VTEQQKSYSKTYY  470
Sp_Tip1p	459	FRKD IWKL CRPF TSMKMEAFAEK INENNVRMD  491
Sc_Est2p	395	FRHD TWNKL TPF VEYFKTYL VENNVCRNHNS  427
Ea_p123	471	YRKNIWDVIMKMSIADLKKE TLAEVQEKEYEEW  503
Sp_Tip1p	492	TQKTTLP PAVIRLL PKK -NTFRL ITNLRKRF  522
Sc_Est2p	428	YTLSNFNH SKM PKSNN EFRI IAIPCRGAD  460
Ea_p123	504	KKSLGFAPGK LRL PKK -TTFRP IMTFNKKIV  534
Sp_Tip1p	523	I KMGSNKKMLVSTINQTLR P VASILKHLINE  552
Sc_Est2p	461	EEE -FTIYKENHKNAI QPTQK ILEYL RNKRPT  491
Ea_p123	535	NSD -RKTTKLTTN TKLLNSHLM LKTNR-MF  564
Sp_Tip1p	553	ESSG PFNLLEVYMKLL TFKKDL KHMFGGR-KK  584
Sc_Est2p	492	SFTKIYSPTQIADR K EFKQRLL KKFNNVLPEEL  524
Ea_p123	565	KDPFGFAVFNYDDV MKYEEFVCKWKQVGQPKL  597
Sp_Tip1p	585	YFVR DIKSCYDBIKQDL MFRIYKKKLKDPE-F  616
Sc_Est2p	525	YFMKF DVKSCYDSIPRMECMR LKDALKNENG  557
Ea_p123	598	FEATMDIEKCYDSVNREKL STFLKTTKLLSSDF  630
Sp_Tip1p	617	VIRKYATIHATSDRATKN --- --- ---  634
Sc_Est2p	558	FVRSQYFFNTNTG --- --- ---  570
Ea_p123	631	WIMTAQILKRKNNIVIDSKNFRKKEMKDYFRQK  663

39/59

FIG. 42  
(CONTINUED)

A.	Sp_Tip1p	635	<b>F</b> V S E A F S Y F D M V P F E K V V Q L L S - - M K T S D T L F V	665
	Sc_Est2p	571	- - - - - V L K L F N V V N A S R - - V P K P Y E L Y -	591
	Ea_p123	664	<b>F</b> Q K I A L E G G Q Y P T L F S V L E N E Q N D L N A K K T L I V	696
Sp_Tip1p	666	<b>D</b> F V D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698	
Sc_Est2p	592	<b>D</b> N V R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624	
Ea_p123	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729	
Sp_Tip1p	699	<b>L</b> Q K V G I P Q G S I L S S F L C H F Y M E D L I D E Y L S F T K	731	
Sc_Est2p	625	I R E D G L F Q G S S L S A P I V D L V Y D D L L E F Y S E F K A	657	
Ea_p123	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762	
Sp_Tip1p	732	K K G - - - - - S V I L L R V V D D F L F I T V N K K D A K K	756	
Sc_Est2p	658	S P S Q D - - - - - T L I L K L A D D F L I I S T D Q Q Q V I N	684	
Ea_p123	763	D E S M N P E N P N V N L L M R L T D D Y L L I T Q E N N A V L	795	
Sp_Tip1p	757	<b>F</b> L N L S L R G F E K H N F S T S L E K T V I N F E N S N G - - -	786	
Sc_Est2p	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - - -	713	
Ea_p123	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828	
Sp_Tip1p	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816	
Sc_Est2p	714	- - - D D T V I Q F C A - - M H I F V K E L E V W K H S S T M	739	
Ea_p123	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861	
Sp_Tip1p	817	<b>A</b> C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849	
Sc_Est2p	740	<b>N</b> N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772	
Ea_p123	862	<b>N</b> I N L R I E G I L C T L N L N M Q T K K A S M W L K K L K S F	894	

*FIG. 42*  
(CONTINUED)

A.								
Sp_Tip1p	850	[L]AS[F]A[Q]V[F]IDIT[H]NSK[F]N[SC]CN YRLGY[SM]CMR	882					
Sc_Est2p	773	[L]NST[N]TV[L]MQ[D]H[V]KN[SEC]	793					
Ea_p123	895	[L]MNN[I]TH[Y]FRKTITTEDFANKTLNKLFISGGYK	927					
Sp_Tip1p	883	AQAY[YL]KRM[KD]IFIPQRMFITD[LL]NVIGRK[WKK]	915					
Sc_Est2p	794	-[Y]KSAF[KD]LSIN-[V]TQNM[QFH]SFLQR[IE]M	821					
Ea_p123	928	YMQCAKE[Y]KDHF[K]KNLAMSSMIDLEVSKIIVYSV	960					
Sp_Tip1p	916	LAEIL[G]YT[SRR]F[LS]SAEV[KWL]FC[G]M[R]D[G]L[K]PS	948					
Sc_Est2p	822	[TV]SGCP[IT]K[C]DPLIEYEVRFTI[LNG]FLES[SS]N	854					
Ea_p123	961	[TRAFF]FKYL[V]CNIKDTIFGEEHYPDFFLSTTLKH[F]	993					
Sp_Tip1p	949	FKYHPCFEQLIYQ[F]QSSLTDL[I]KPLRPVL[R]QVL[F]	981					
Sc_Est2p	855	TS-[K]FDNI[IL]LRKEIQHLQAYIY	877					
Ea_p123	994	IEIFS-[T]KKYIF[N]RVCM[IL]KAKEAKL[K]SDQC	1023					
Sp_Tip1p	982	LHRRRIAD-	988					
Sc_Est2p	878	IYIHIVN-	884					
Ea_p123	1024	QSLI[Q]YDA	1031					

FIG. 42  
(CONTINUED)

B.

Sp_Tip1p	1	- - - - -	- M T E H H T P K S R I L R F L E N Q V Y V L C T	24
Sc_Est2p	1	- - - - -	- - - - - M K I L F E F	7
Ea_p123	1	M E V D V D N Q A D N H G I H S A L K T C E E I K E A K T L Y S W	33	
Sp_Tip1p	25	L N D Y V Q L V L R G S S P A [S Y S N I C E R L R S D V Q T S F S	57	
Sc_Est2p	8	I Q D K L D I D L Q T N - - S T Y K - - E N L K C G H F N G L D	35	
Ea_p123	34	I Q K V I R C R N Q S Q - - S H Y K - - D L E D I K I F A Q T N	61	
Sp_Tip1p	58	I F L H S T V V G F D S K P D E G V Q F F S S P K C S Q S E L [I A N	90	
Sc_Est2p	36	E I L T C F A L P N S R - K I A L P C L P G D L S H K A V [D H	67	
Ea_p123	62	I V A T P R D Y N E E D F K V I A R K E V F S T G L M I E L [D K	94	
Sp_Tip1p	91	V V K Q M F D E S F E R R R - N L L M K [G F S M N H E D F R A M H	122	
Sc_Est2p	68	C I Y L L T G E L Y N - - N V I T F G Y K I A R N E D - - -	93	
Ea_p123	95	C L V E L L S S D V S D R Q K L Q C F G F Q L K G N Q - - -	122	
Sp_Tip1p	123	V N G V Q N D L V S T F P N Y L I S I L E S K N [W Q L L E I I G	155	
Sc_Est2p	94	- - - V N N S L F C H S A N V N V T L L K G A A W K M F H S L V G	123	
Ea_p123	123	- - - L A K T H L L T A L S T Q K Q Y F F Q D E W N Q V R A M I G	152	
Sp_Tip1p	156	S D A M H Y [L S K G S I F E A L P N D N Y L Q I S G I P L F K N	188	
Sc_Est2p	124	T Y A F V D L L I N Y T V I Q F N - G Q F F T Q I V G N R C N E P	155	
Ea_p123	153	N E L F R H L Y T K Y L I F Q R T S E G T L V Q F C G N N V F D H	185	
Sp_Tip1p	189	N V F E E T V S K K R K R T I E T S I T Q N - - - K S A R K E V S	218	
Sc_Est2p	156	H L P P K W V Q - - R S S S S A T A A Q I - - - K Q L T E P V T	183	
Ea_p123	186	L K V N D K F D K - K Q K G G A A D M N E P R C C S T C K Y N V K	217	

FIG. 42  
(CONTINUED)

+

42/59

+

## B.

<b>Sp_Tip1p</b>	219	WNSISISRFSSIFYRSSYKKFKQDLYFNLHSICD	251
<b>Sc_Est2p</b>	184	N-----KQFLHKLNINSSSFFP	200
<b>Ea_p123</b>	218	NEK-----FYCTHFN	248
<b>Sp_Tip1p</b>	252	RNTVHMWLQWIFPRQFGLINAFAQVKQLHKGIVPL	284
<b>Sc_Est2p</b>	201	-----YSKILPSSS-----SIKKLTDLREAIFP	223
<b>Ea_p123</b>	249	R-----NNQFFFFKHEFVSNNKNNISAMDRAQTI	275
<b>Sp_Tip1p</b>	285	VS-----QSTVVPKRLLKVVYPPLIEQTAKRILHRS	313
<b>Sc_Est2p</b>	224	TN-----LVKIPQRLLKVRINLTLQKLLKRHKRLN	252
<b>Ea_p123</b>	276	FTNIFRFNRIRKLLKDKVIEKIAYMLEVKDFN	308
<b>Sp_Tip1p</b>	314	LSKVYNHYCPYIDTHDDEKILSYSLKPNQ-----	342
<b>Sc_Est2p</b>	253	YVSILNSICPPLEGTVLDLSHLSRQSPKER-----	282
<b>Ea_p123</b>	309	FNYYLTKSCPLPENWRERKQQKENLINKTREEK	341
<b>Sp_Tip1p</b>	343	-----VFAFLRSILVRYVFPKL-----	359
<b>Sc_Est2p</b>	283	-----VLKFIIVILQKLLPQEM-----	299
<b>Ea_p123</b>	342	SKYYEELFSYTTDNKC-----VTPQFINEFFYNILPKDF	374
<b>Sp_Tip1p</b>	360	WGNGQRIFEILKDELTFLKLSRYESSFLHYLMS	392
<b>Sc_Est2p</b>	300	FGSKKNKGKIKNLNLLSLPLNGYLPFDSSLK	332
<b>Ea_p123</b>	375	LTG-RNRKNFQKKVKKYVELNKHELISHKNLLE	406
<b>Sp_Tip1p</b>	393	NIKISEIEWLVLGKR SNAKMCLSDFEKRKQIFA	425
<b>Sc_Est2p</b>	333	KLRLKDFRWLFI-----DIWFTKHNFENLNQLAI	362
<b>Ea_p123</b>	407	KINTREISWMQVETS-AKHFFYYFDHEN-IYVLW	437

B.		
Sp_Tip1p	426	E F I Y W L Y N S F I I P I L Q S [F F Y I T E S S D L R N R T V Y
Sc_Est2p	363	C F I S W L F R Q L I P K I Q T F F Y C T E I S S T V T - I V Y
Ea_p123	438	K I L R W I F E D L V V S L I R C F F Y V T E Q Q K S Y S K T Y Y
Sp_Tip1p	459	F R K D I W K L L C R P F I T S M K M E A F E K I N E N N V R M D
Sc_Est2p	395	F R H D T W N K L I T P F I V E Y F K T Y L V E N N V C R N H N S
Ea_p123	471	Y R K N I W D V I M K M S I A D L K K E T L A E V Q E K E V E E W
Sp_Tip1p	492	T Q K T T L P P A V I R L L P K K - - N T F R L I T N L R K R F L
Sc_Est2p	428	Y T L S N F N H S K M R I I P K K S N N E F R I I A I P C R G A D
Ea_p123	504	K K S L G F A P G K L R L I P K K - - T T F R P I M T F N K K I V
Sp_Tip1p	523	I K M G S N K K M L V S T N Q T L R P V A S I L K H L I N E - - -
Sc_Est2p	461	E E E - - F T I Y K E N H K N A I Q P T Q K I L E Y L R N K R P T
Ea_p123	535	N S D - - R K T T K L T T N T K L L N S H L M L K T L K N R - M F
Sp_Tip1p	553	E S S G I P F N L E V Y M K L L T F K K D L L K H R M F G R - K K
Sc_Est2p	492	S F T K I Y S P T Q I A D R I K E F K Q R L L K K F N N V L P E L
Ea_p123	565	K D P F G F A V F N Y D D V M K K Y E E F V C K W K Q V G Q Q P K L
Sp_Tip1p	585	Y F V R I D I K S C Y D R I K Q D L L M F R I V K K K L K D P E - F
Sc_Est2p	525	Y F M K F D V K S C Y D S I P R M E C M R I L K D A L K N E N G F
Ea_p123	598	F F A T M D I E K C Y D S V N R E K L S T F L K T T K L L S S D F
Sp_Tip1p	617	V I R K Y A T I H A T S D R A T K N - - - - -
Sc_Est2p	558	F V R S Q Y F F N T N T G - - - - -
Ea_p123	631	W I M T A Q I L K R K N N I V I D S K N F R K K E M K D Y F R Q K

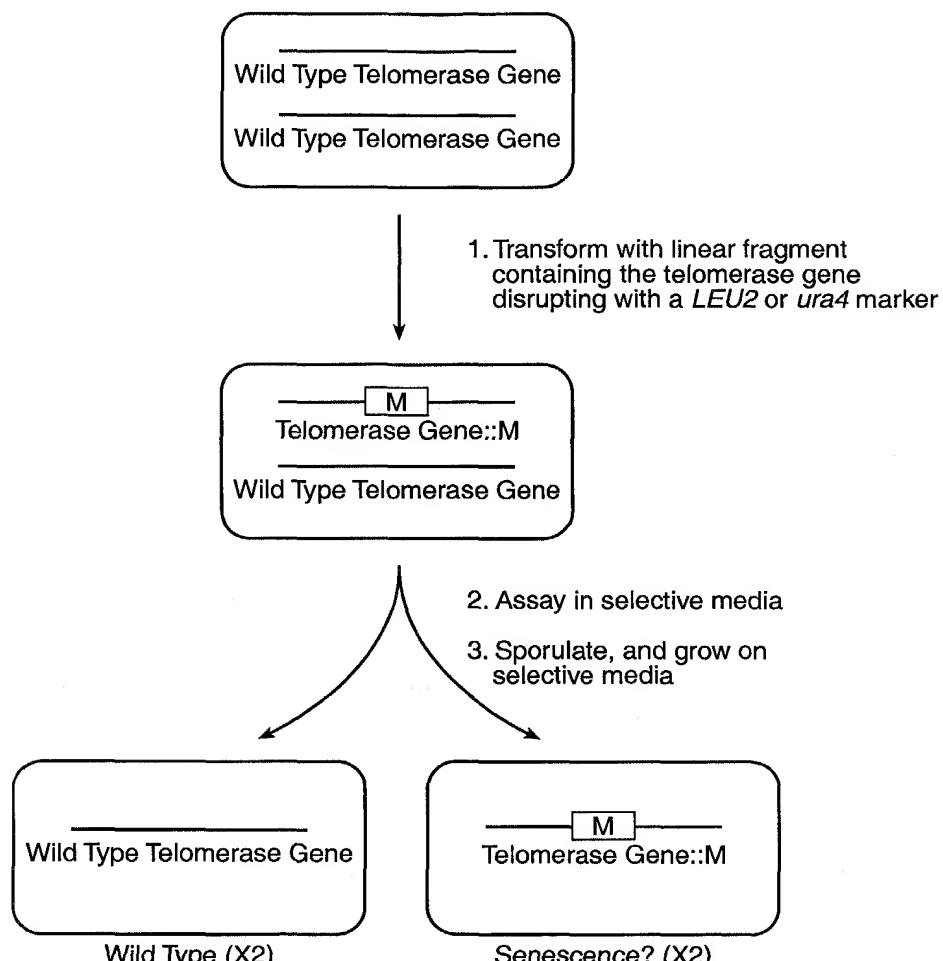
B.

<b>Sp_Tip1p</b>	635	F V S E A F S Y F D M V P F E K [V] V Q L L S - - M K T S D T L [F] V	665
<b>Sc_Est2p</b>	571	- - - - - V L K L F N V V N A S R - - V P K P Y E L [Y] I	591
<b>Ea_p123</b>	664	F Q K I A L E G G Q Y P T L F S [V] L E N E Q N D L N A K K T L I V	696
<b>Sp_Tip1p</b>	666	D F V D Y W T K S S S E I F K M L K E H L S G H I V K I G N S Q Y	698
<b>Sc_Est2p</b>	592	D N V R T V H L S N Q D V I N V V E M E I F K T A L W V E D K C Y	624
<b>Ea_p123</b>	697	E A K Q R N Y F K K D N L L Q P V I N I C Q Y N Y I N F N G K F Y	729
<b>Sp_Tip1p</b>	699	L Q K V G I P Q G S I L [S] S F L C H F Y M E D L I D E Y L S F T K	731
<b>Sc_Est2p</b>	625	I R E D G L F Q G S S L [S] A P I V D L V Y D D L L E F Y S E F K A	657
<b>Ea_p123</b>	730	K Q T K G I P Q G L C V S S I L S S F Y Y A T L E E S S L G F L R	762
<b>Sp_Tip1p</b>	732	K K G - - - - S V L L R V V D D F L F I T V N K D A K K	756
<b>Sc_Est2p</b>	658	S P S Q D - - - - T L I L K L A D D F L I I S T D Q Q Q V I N	684
<b>Ea_p123</b>	763	D E S M N P E N P V N L L M R L T D D Y L L I T T Q E N N A V L	795
<b>Sp_Tip1p</b>	757	F L N L S L R G F E K H N F S T S L E K T V I N F E N S N G - -	786
<b>Sc_Est2p</b>	685	I K K L A M G G F Q K Y N A K A N R D K I L A V S S Q S D - -	713
<b>Ea_p123</b>	796	F I E K L I N V S R E N G F K F N M K K L Q T S F P L S P S K F A	828
<b>Sp_Tip1p</b>	787	- - - I I N N T F F N E S K K R M P F F G F S V N M R S L D T L L	816
<b>Sc_Est2p</b>	714	- - - D D T V I Q F C A - - M H I F V K E L E V W K H S S T M	739
<b>Ea_p123</b>	829	K Y G M D S V E E Q N I V Q D Y C D W I G I S I D M K T L A L M P	861
<b>Sp_Tip1p</b>	817	A C P K I D E A L F N S T S V E L T K H M G K S F F Y K I L R S S	849
<b>Sc_Est2p</b>	740	N N F H I R S K S S K G I F R S L I A L F N T R I S Y K T I D T N	772
<b>Ea_p123</b>	862	N I N L R I E G I L C T L N L N M Q T K K A S M W L K K K L K S F	894

B.							
Sp_Tip1p	850	[LASFAQVFIDITHNSKFNNSCCNIVRLGYSMCMR	882				
Sc_Es12p	773	LNSTNTVLMQIDHVVKNISEC-	793				
Ea_p123	895	LMNNITHYFRKTITTEDFANKTLNKLFISGGYK	927				
Sp_Tip1p	883	AQAYLKRMDIFIPQRMFITDLLNVIGRKIWKK	915				
Sc_Es12p	794	- - YKSAFKDLSIN- - VTQNMQFHSFLQRIIEM	821				
Ea_p123	928	YMQCAKEYKDHFKKNLAMSSMIDLEVSKIIYSV	960				
Sp_Tip1p	916	LAEILGYTSRRFLSSAEVKWLFCLGMRDGGLKPS	948				
Sc_Es12p	822	TVSGCPITKCDPLIEYEVRFTILNGFLESLSNN	854				
Ea_p123	961	TRAFFKYLVVCNIKDTIFGEEHYPDFFFLSTLKHF	993				
Sp_Tip1p	949	FKYHPCFEQLIYQFQSQLTDLIKPLRPVLRQVLF	981				
Sc_Es12p	855	TS- - - - KFKDNILLRKEIQHLQAYIY	877				
Ea_p123	994	IEIFS- - - TKKYIFENRVCMILKAKEAKLKSDDQC	1023				
Sp_Tip1p	982	LHRRIAD-	988				
Sc_Es12p	878	IYIHIVN-	884				
Ea_p123	1024	QSLIQYDA	1031				

FIG. 42  
(CONTINUED)

P0232250



(These cells will show a senescence phenotype  
if the disrupted gene encodes a telomerase subunit.)

FIG. 43

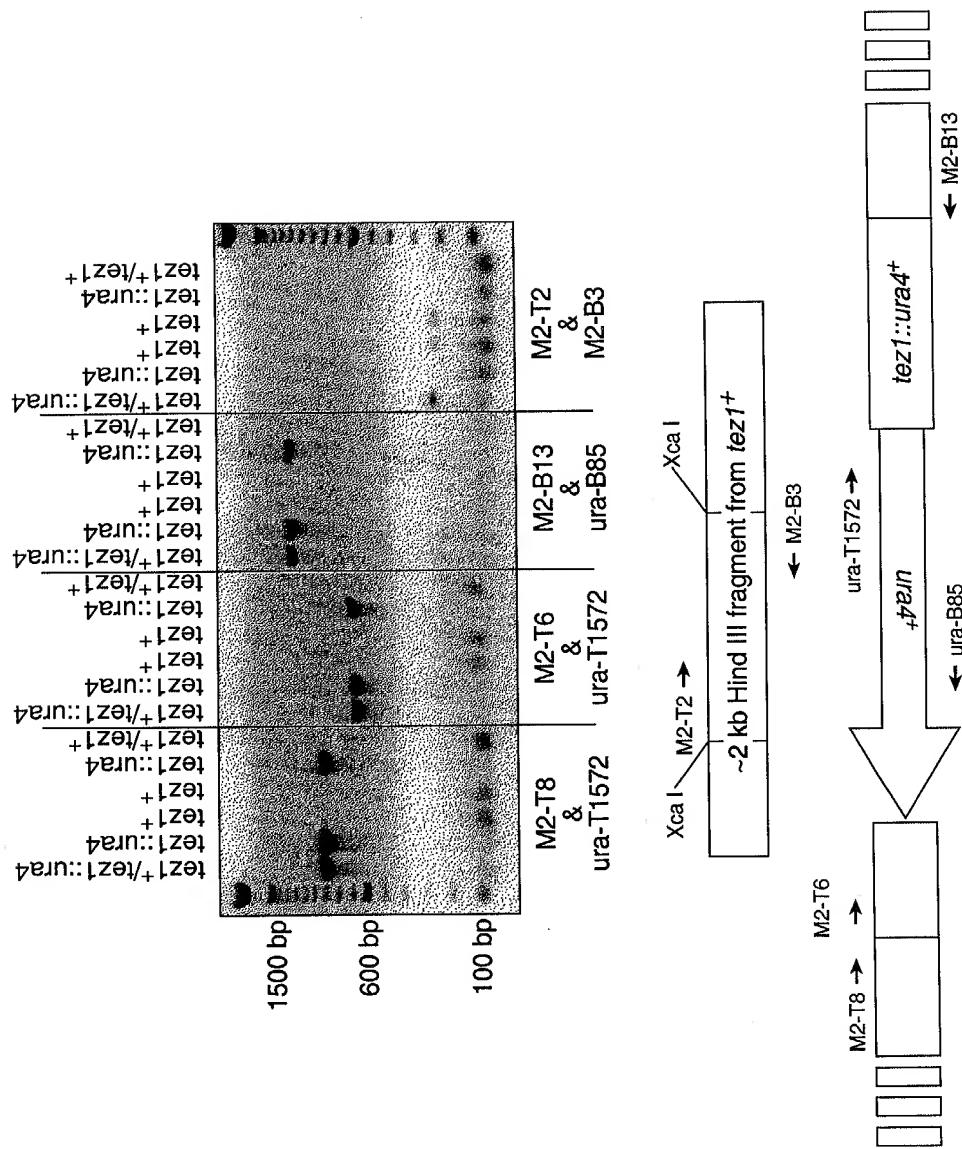


FIG. 44

+

49/59

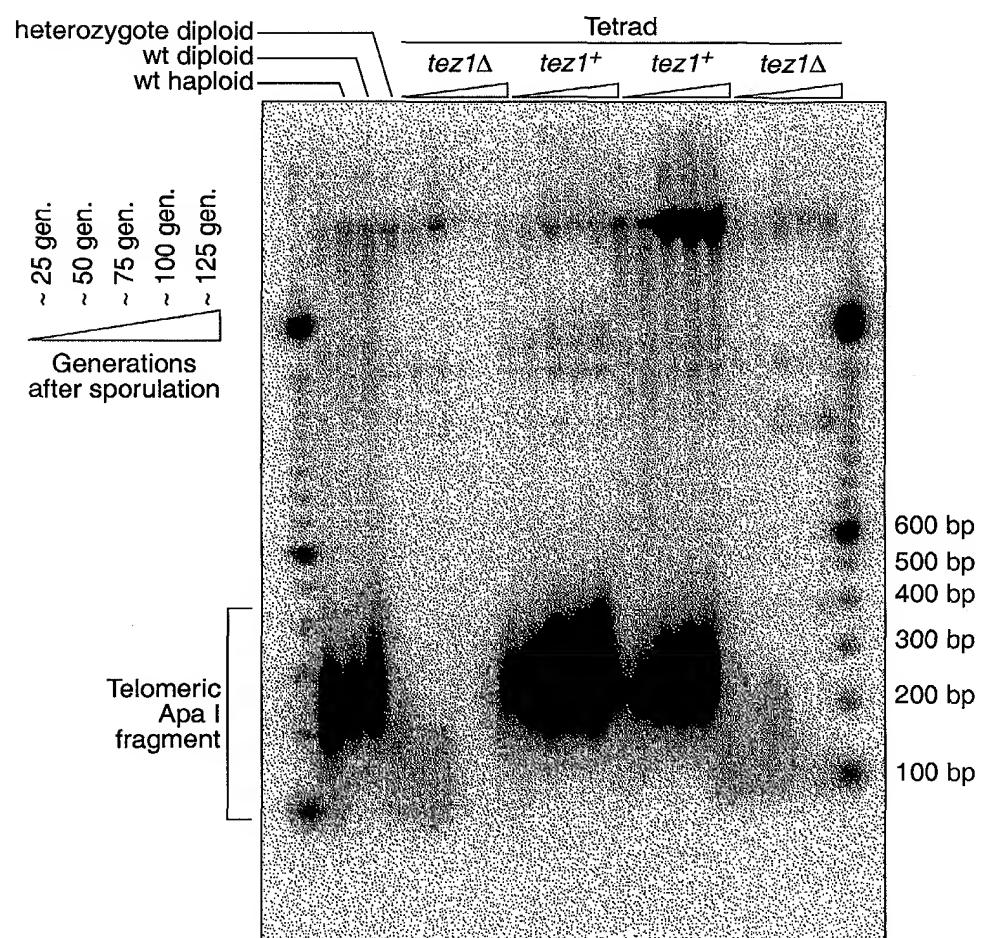


FIG. 45

+

FIG. 46

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT GAG TCA AAA AAT TGG CAA 1529  
 129 D L V S T F P N Y L I S I L E S K N W Q 148  
 1530 CTT TTG TTA GAA AT gtaaaataccggtaagatgttgcgcactttgacaaggatatacgat 1601  
 149 L L L E I I G 155  
 1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661  
 156 S D A M H Y L L S K G S I F E A L P N D 175  
 1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721  
 176 N Y L Q I S G I P L F K N N V F E E T V 195  
 1722 TCA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCG CGC AAA 1781  
 196 S K K R K R T I E T S I T Q N K S A R K 215  
 1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841  
 216 E V S W N S I S R F S I F Y R S S Y 235  
 1842 AAG AAG TTT AAG CAA G gtaactaatactgttatccatataacttaatttttag AT CTA TAT TTT AAC 1907  
 236 K K F K Q D L Y F N 245  
 1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967  
 246 L H S I C D R N T V H M W L Q W I F P R 265  
 1968 CAA TTT GGA CTT ATA AAC GCA ATT CAA GTG AAG CAA TTT CAC AAA GTG ATT CCA CTG GTA 2027  
 266 Q F G L I N A F Q V K Q L H K V I P L V 285  
 2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087  
 286 S Q S T V V P K R L L K V Y P L I E Q T 305  
 2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147  
 306 A K R L H R I S L S K V Y N H Y C P Y I 325  
 2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT ATT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207  
 326 D T H D D E K I L S Y S L K P N Q V F A 345  
 2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267  
 346 F L R S I L V R F P K L I W G N Q R I 365

**FIG. 46**  
 (CONTINUED)

2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaatttaccactaacgattttaccag AC CTC GAA ACT 2336  
 366 F E I I L K D  
 2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396  
 376 F L K L S R Y E S F S L H Y L M S N I K 395  
 2397 gtaatatgccaaattttaccattaataacatcg ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465  
 396 I S E I E W L V L G 405  
 2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525  
 406 K R S N A K M C I S D F E K R Q I F A 425  
 2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585  
 426 E F I Y W L Y N S F I I P I L Q S F F Y 445  
 2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645  
 446 I T E S D L R N R T V Y F R K D I W K 465  
 2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705  
 466 L C R P F I T S M K M E A F E K I N E 485  
 2706 gtatttaaggatttttggaaaaaggctaattttcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775  
 486 N N V R M D T Q K T 495  
 2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAT ACC TTT CGT CTC ATT ACG 2835  
 496 T L P P A V I R L L P K K N T F R L I T 515  
 2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaaatttttgtcattcaatgtactttacttctaattata 2906  
 516 N L R K R F L I K 524  
 2907 tttagcag ATG GGT TCA AAC AAA ATG TTA GTC AGT AAC CAA ACT TTA CGA CCT GTG 2967  
 525 M G S N K K M L V S T N Q T L R P V 542  
 2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027  
 543 A S I L K H L I N E E S S G I P F N L E 562  
 3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT AAG CAC CGA ATG TTT GG gtaat 3088  
 563 V Y M K L L T F K K D L L K H R M F G 581

**FIG. 46**  
(CONTINUED)

3089 tataataatgcgcgattccctcatttgcag G CGT AAG TAT TTT GTA CGG ATA GAT ATA  
 R K Y F V R I D I 3155  
 582 S C Y D R I K Q D L M F R I V K K L 591

3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC  
 592 K S C Y D R I K Q D L M F R I V K K L 3215  
 611

3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT  
 612 K D P E F V I R K Y A T I H A T S D R A 3275  
 631

3276 ACA AAA AAC TTT GTT ACT GAG GCG TTT TCC TAT T gtaagtttatttttaattttaacaa 3343  
 632 T R N F V S E A F S Y F 643

3344 attcttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA  
 644 D M V P F E K V V Q L L S M K T 3405  
 659

3406 TCA GAT ACT TTG TTT GAT TTT GTG GAT TAT TTG ACC AAA AGT TCT TCT GAA ATT TT  
 660 S D T L F V D F V D Y W T K S S E I F 3465  
 679

3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaaatgttaataaca 3532  
 680 K M L K E H L S G H I V K 692

3533 ctaatggaaacttag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA  
 693 I G N S Q Y L Q K V G I P Q G S 3593  
 708

3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG  
 709 I L S S F L C H F Y M E D L I D E Y L S 3653  
 728

3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA  
 729 F T K K K G S V L L R V D D F L F I T 3713  
 748

3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgtgtcattcc 3777  
 749 V N K K D A K K F L N L S L R G 764

3778 taaggttcaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA  
 765 F E K H N F S T S L E K T V 3840  
 778

3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT ATT GAA ACC AAG AAA  
 779 I N F E N S N G I I N N T F F N E S K K 3900  
 798

**FIG. 46**  
(CONTINUED)

54/59

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960  
799 R M P F F G F S V N M R S L D T L L A C 818

3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTC GAG CTG ACG AAA CAT ATG GGG 4020  
819 P K I D E A L F N S T S V E L T K H M G 838

4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaaataataggtaatcg A TCG 4089  
839 K S F F Y K I L R S 848

4090 AGC CTT GCA TCC TTT GCA CAA GTC ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149  
849 S L A S F A Q V F I D I T H N S K F N S 868

4150 TGC TGC AAT ATA TAT AGG CTA GCA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209  
869 C C N I Y R L G Y S M C M R A Q A Y L K 888

4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtggatcttttaactaga 4274  
889 R M K D I F I P Q R M F I T D 903

4275 aaagtccattaaatttaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AGG TTG GCC 4339  
904 L L N V I G R K I W K K L A 917

4340 GAA ATA TTA CGA TAT AGC AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401  
918 E I L G Y T S R R F L S S A E V K W 935

4402 ggtctcgagacttcagcaataattgcacatcg G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468  
936 L F C L G M R D G L K 946

4469 CCC TCT TTC AAA TAT GAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528  
947 P S F K Y H P C F E Q L I Y Q F Q S L T 966

4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA AGA ATA 4588  
967 D L I K P L R P V L R Q V L F L H R R I 986

4589 GCT GAT TAA tgtcatttcaatttattatacatcccttattactgggtcttaacaattattactaaggata 4665  
987 A D \* 989

FIG. 46  
(CONTINUED)

*FIG. 46*  
(CONTINUED)

1  
met ser val tyr val val glu leu leu  
GCCAAGTTCCCTGCACTGGCTG ATG AGT GTG TAC GTC GTC GAG CTG CTC

10  
arg ser phe phe tyr val thr glu thr thr phe gln lys asn arg  
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

20  
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile  
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

30  
gly ile arg gln his leu lys arg val gln leu arg glu leu ser  
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

40  
glu ala glu val arg gln his leu lys arg val gln leu arg glu leu ser  
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

50  
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro  
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

60  
ile val asn met asp tyr val val gly ala arg thr phe arg arg  
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

70  
glu lys ala glu arg leu thr ser arg val lys ala leu phe  
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

80  
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly  
AGC GTG CTC AAC TAC GAG CGG GCG CGC CCC GGC CTC CTG GGC

90  
ala ser val leu gly leu asp asp ile his arg ala trp arg thr  
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

100  
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr  
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG TAC

110  
phe val lys val asp val thr gly ala tyr asp thr ile pro gln  
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

120  
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn  
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC

130  
thr tyr cys val arg arg tyr ala val val gln lys ala ala met  
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

140

150

160  
170  
180  
190  
200

FIG. 47

210

gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys  
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220

gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser  
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

230

240

leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg  
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250

arg asp gly leu leu leu arg leu val asp asp phe leu leu val  
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG GTG

260

270

thr pro his leu thr his ala lys thr phe leu arg thr leu val  
ACA CCT CAC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280

arg gly val pro glu tyr gly cys val val asn leu arg lys thr  
CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG AAG ACA

290

300

val val asn phe pro val glu asp glu ala leu gly gly thr ala  
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310

320

phe val gln met pro ala his gly leu phe pro trp cys gly leu  
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

330

340

leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser  
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

350

tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly  
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

360

phe lys ala gly arg asn met arg arg lys leu phe gly val leu  
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370

380

arg leu lys cys his ser leu phe leu asp leu gln val asn ser  
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

390

400

leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln  
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

410

ala tyr arg phe his ala cys val leu gln leu pro phe his gln  
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

**FIG. 47**  
(CONTINUED)

+

58/59

420

gln val trp lys asn pro his phe ser cys ala ser ser leu thr  
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430

440

arg leu pro leu leu leu his pro glu ser gln glu arg arg asp  
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450

val ala gly gly gln gly arg arg arg pro ser ala leu arg gly  
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460

470

arg ala val ala val pro pro ser ile pro ala gln ala asp ser  
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480

thr pro cys his leu arg ala thr pro gly val thr gln asp ser  
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490

500

pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys  
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

510

pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp  
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520

530

his pro gly leu met ala thr arg pro gln pro gly arg glu gln  
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540

thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly  
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550

560

arg gly gly pro his pro gly leu his arg trp glu ser glu ala  
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564

OP

TGA GTGAGTGTTGGCCGAGGCCCTGCATGTCCGGCTGAAGGCTGAGTGTCGGCTGAGGC  
CTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTCACTCCCCAC  
AGGCTGGCGTTCGGTCCACCCCAGGGCCAGCTTCTCACCAAGGAGCCGGCTTCCACT  
CCCCACATAGGAATAGTCATCCCCAGATTGCCATTGTTCACCCCTGCCCTGCCTTCC  
TTTGCCTTCCACCCCCACCATTAGGTGGAGACCCTGAGAAGGACCCCTGGAGCTTGGG  
AATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGGCAGGACCCCTGCACCTGGATGGGG  
GTCCCTGTGGGTCAAATTGGGGGAGGTGCTGTGGAGTAAAATACTGAATATATGAGTT  
TTTCAGTTTGGAAAAAAAAAAAAAA

FIG. 47  
(CONTINUED)

Motif -1						
Ep p123	...LVVSLIRCFYYVTEQQKSYSKT...					
Sp Tez1	...FIIPILQSFFYITESSDLRNRT...					
Sc Est2	...LIPKIIQTFFYCTEISSTVTIV...					
Hs TCP1	...YVVELLRSFFYVTEFFQKNRL...					
consensus	FFY TE					
 Motif 0						K
Ep p123	p hhh K hR h R					
Sp Tez1	...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...					
Sc Est2	...QKTTLPPAVIRLLPKKN--TFRPLITNLRKRLF...					
Hs TCP1	...TLSNFNHSKMRIIPKKSNNEFRIIAIPCRGAD...					
consensus	...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...					
	R PK	R I				
 Motif A						AF
Ep p123	h hDh GY h					
Sp Tez1	...PKLFFATMDIEKCYDSVNREKLSTFLK...					
Sc Est2	...RKKYFVRIDIJKSCYDRIKQDLMFRIVK...					
Hs TCP1	...PELYFMKFDVKSCYDSIPRM ECMRILK...					
consensus	...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...					
	F D YD					
 Motif B						hPQG pS hh
Ep p123	...NGKFYKQTKGIPQGLCVSSILSSFYYA...					
Sp Tez1	...GNSQYLQKV GIPQGSILSSFLCHFYME...					
Sc Est2	...EDKCYIREDGLFQGSSL SAPIVDLVYD...					
Hs TCP1	...RATSYVQCQGIPQGSILSTLLCSLCYG...					
consensus	G QG S					
 Motif C						Y
Ep p123	h F DD hhh					
Sp Tez1	...PNVNLLMRLTDYLLIT TQENN...					
Sc Est2	...KKGSVLLRVVDDFLFITVNKKD...					
Hs TCP1	...SQDTLILKLADDFLIIISTDQQQ...					
consensus	...RRDGLLLRLVDDFLLVTPHLTH...					
	DD L					
 Motif D						Gh h cK
Ep p123	...NVSRENGFKFNMKKL...					
Sp Tez1	...LNLSLRGF EKHN FST...					
Sc Est2	...KKLAMGGFQKYNAKA...					
Hs TCP1	...LRTLVRGVPEYGCVV...					
consensus	G					

FIG. 48